

Numerische Verfahren zur Identifizierung von Fehlstellen aus Randdaten

von

Marek Engelhardt

Institut für Angewandte Mechanik

Technische Universität Braunschweig

Herausgegeben vom Mechanik-Zentrum der
Technischen Universität Braunschweig

Schriftleiter: Prof. Dr.rer.nat. H. Antes
Institut für Angewandte Mechanik
Postfach 3329
38023 Braunschweig

Vom Fachbereich Bauingenieurwesen der Technischen Universität
Carolo-Wilhelmina zu Braunschweig zur Erlangung des Grades eines
Doktor-Ingenieur (Dr.-Ing.) genehmigte Dissertation

Tag der Einreichung: 05.05.2004
Tag der Prüfung: 02.07.2004

Berichter:

Prof. Dr.-Ing. G.E. Stavroulakis
Prof. Dr.rer.nat. H. Antes
Prof. Dr.-Ing. K.-U. Bletzinger

© Copyright 2004 Marek Engelhardt, Braunschweig

BSM 56-2004
ISBN 3-920395-55-7

Alle Rechte, insbesondere der Übersetzung in fremde Sprachen, vorbehalten.
Mit Genehmigung des Autors ist es gestattet, dieses Heft ganz oder teilweise
zu vervielfältigen.

Zusammenfassung

Das Erkennen und nachfolgend die Lokalisierung einer Schädigung eines Bauteils mittels gemessener charakteristischer Daten führt auf ein inverses Problem, das in dieser Arbeit als Optimierungsaufgabe formuliert und anschließend numerisch gelöst wird. Das mechanische Verhalten des Bauteils wird mit der Randelementmethode nachgebildet, wobei spezielle Elementtypen zur Simulierung von Rissen innerhalb des Bauteils eingesetzt sowie Untersuchungen mit statischen und dynamischen Belastungen durchgeführt werden. Zur Identifizierung der Schädigungen werden Optimierungsstrategien mit Algorithmen aus der Neuroinformatik angewandt. Es zeigt sich, daß die Identifikation unterschiedlicher Schädigungen mittels numerischer Verfahren sowohl in der Statik als auch in der Dynamik erfolgreich durchgeführt werden kann, wobei nicht nur unterschiedliche Geometrien und die Anzahl der Defekte, sondern auch die Genauigkeit der Messungen einen deutlichen Einfluß auf die Identifikation haben. Die Belastungsart und die Anzahl der Lastfälle sind dabei auf die Geometrie abzustimmen.

Genetische Algorithmen sind sehr gut zur Schadenserkenkung in statisch belasteten Strukturen geeignet, allerdings sehr rechenaufwendig. Die Ergebnisse bestätigen die Robustheit und die Zuverlässigkeit dieser Algorithmen. Wegen ihres hohen Rechenaufwands können diese Verfahren noch nicht auf dynamisch belastete Bauteile angewandt werden.

Als auch bei dynamisch belasteten Strukturen geeignete Verfahren zur Fehlstellenidentifikation zeigen sich die Methoden des Kalman-Filters sowie der neuronalen Netze.

Abstract

Material defects, such as cracks and flaws, within structural components are identified and localized from non-destructive testing data. In this study, the resulting inverse problem is reformulated as an optimization problem and solved numerically. The mechanical behaviour of a component is simulated using the boundary element method. Special elements for modelling cracks are applied. Optimization strategies including soft computing algorithms are used to analyze the influence of defects on boundary reactions caused by static as well as dynamic loadings.

Applying these methodologies, various defects can be identified in both the static and dynamic setting. The geometry of the considered structural members determines the nature and number of the applied loadings. Not only the type and number of defects, but also measurement errors have a significant influence on the identification results.

In statically loaded structures, genetic algorithms are well suited to the detection of defects. The results confirm the robustness and reliability of these algorithms. Unfortunately, they are also very computationally expensive. Hence, in most cases, they cannot be applied to analyze components which are loaded dynamically.

In dynamically loaded structures, Kalman filter techniques as well as neural networks can be applied to efficiently identify cracks.

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Angewandte Mechanik der Carolo-Wilhelmina Universität zu Braunschweig und am Institut der Mathematik der Universität Ioannina in Griechenland.

Allen, die zu Ihrem Gelingen beigetragen haben, möchte ich an dieser Stelle herzlich danken.

An erster Stelle möchte ich mich bei meinem Lehrer und Doktorvater Prof. Dr.-Ing. G. E. Stavroulakis für die Anregung zu der untersuchten Thematik bedanken.

Weit über die alleinige Übernahme des Hauptreferates hinaus war Prof. Dr.-Ing. G. E. Stavroulakis durch sein großes Interesse, seine stetige Diskussionsbereitschaft und seine Offenheit gegenüber neuen Ideen eine wesentliche Unterstützung bei der Anfertigung der Dissertation.

Mein besonderer Dank gilt auch Prof. Dr.rer.nat. H. Antes, der mir als Koreferent in ausführlichen Diskussionen auch über die Arbeit hinausgehend wertvolle Hinweise geben konnte und somit einen unverzichtbaren Beitrag zum Entstehen der Arbeit gab.

Ebenso möchte ich mich bei PD Dr.-Ing. M. Schanz für seine stetige Diskussionsbereitschaft und für die sorgfältige Durchsicht dieser Arbeit bedanken.

Für die Übernahme des Prüfungsvorsitzes möchte ich mich bei Prof. Dr.-Ing. T. Ummenhofer bedanken. Ebenso bedanke ich mich bei Prof. Dr.-Ing. K.-U. Bletzinger für die kurzfristige Übernahme der Begutachtung der Arbeit als dritter Gutachter.

Mein Interesse und die Freude an der wissenschaftlichen Arbeit resultiert mit Sicherheit zu einem Großteil aus den hervorragenden Rahmenbedingungen, die mir an beiden Instituten geboten wurden, und für deren Existenz ich mich bei Prof. Dr.rer.nat. H. Antes und Prof. Dr.-Ing. G. E. Stavroulakis bedanken möchte.

Fachübergreifend möchte ich mich bei meinen Arbeitskollegen am Institut der Angewandten Mechanik in Braunschweig als auch am Institut der Mathematik in Ioannina für die angenehme Arbeitsatmosphäre bedanken. Es hat mich sehr gefreut, dass auch über die Arbeit hinaus neue Freundschaften aufgebaut werden konnten.

Dank gilt allen an der Arbeit beteiligten wissenschaftlichen Hilfskräften, die mir die Arbeit durch unzählige numerische Studien zur Rissanalyse erleichtert haben.

Darüber hinaus möchte ich mich bei meinen Eltern, Lisa und Harald Engelhardt, meiner Schwester Isabell und meinen Großeltern bedanken, die durch ihre jahrelange Unterstützung meiner Ausbildung und das beständige Interesse an meiner Arbeit zum Gelingen der Promotion beigetragen haben.

Zum Abschluß, aber damit umso herzlicher, möchte ich mich bei meiner Freundin Imke bedanken, die mich immer unterstützt hat und für das notwendige private Umfeld gesorgt hat, das zum Abschalten von der Arbeit notwendig war.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Stand der Forschung	1
1.2	Ziel und Inhalt der Arbeit	3
2	Theoretische Grundlagen	5
2.1	Grundgleichungen der Elastizitätstheorie	5
2.2	Zeitabhängige Randintegralgleichung	6
2.2.1	Fundamentallösung	6
2.2.2	Herleitung mit der Methode der gewichteten Residuen	7
2.2.3	Randintegralgleichung	8
2.2.4	Hypersinguläre Randintegralgleichung	10
2.2.5	Randelementformulierung	12
2.3	Kontaktmechanik	13
2.3.1	Formulierung des Kontaktes	13
2.3.2	Lemke-Verfahren	15
2.4	Stochastische Grundlagen	17
2.4.1	Skalare Zufallsvariablen	18
2.4.2	Vektorielle Zufallsvariablen	22
3	Optimierungsmethoden	26
3.1	SQP - Sequential Quadratic Programming	26
3.2	Kalman-Filter	28
3.2.1	Stochastischer Signalprozeß	28
3.2.2	Lineare Kalman-Filterung in diskreter Zeit	30
3.2.3	Regulärer Kalman-Filteralgorithmus	34
3.2.4	Nichtlineare Kalman-Filterung	38

3.3	Genetischer Algorithmus	42
3.4	Neuronales Netz	45
4	Loch- und Rißerkennung unter statischer Belastung	49
4.1	Identifikation mit SQP	50
4.2	Identifikation mit dem linearen Kalman-Filter	55
4.3	Identifikation mit Neuronalen Netzen	59
4.3.1	Einfluß unterschiedlicher Lernstrategien	59
4.3.2	Identifikation einer kreisförmigen Fehlstelle	65
4.3.3	Identifikation von ellipsenförmigen Fehlstellen	67
4.4	Identifikation mit dem genetischen Algorithmus	72
4.4.1	Identifikation von elliptischen Schädigungen mit dem genetischen Algorithmus in einer Scheibe	73
4.4.2	Identifikation von elliptischen Schädigungen mit einem genetischen Algorithmus im Dreidimensionalen	78
5	Loch- und Rißerkennung im Zeitbereich	98
5.1	Identifikation mit dem linearisierten Kalman-Filter	98
5.1.1	Modellierung der Struktur	99
5.1.2	Anpassung des Kalman-Filters	101
5.1.3	Identifikation einer Schädigung	102
5.1.4	Einfluß der Schädigungsform	104
5.1.5	Unterschiedliche Belastungsarten	109
5.1.6	Schädigungsidentifikation mit verrauschten Meßgrößen	121
5.1.7	Identifikation mehrerer Schädigungen	123
5.1.8	Einfluß eines zweiten Lastfalls	123
5.1.9	Einfluß unterschiedlicher Schädigungsformen	126
5.1.10	Identifikation von vier Rissen	128
5.1.11	Schädigungsidentifikation in nicht rechteckigen Scheiben	129
5.1.12	L-förmige Scheibe	129
5.1.13	U-förmige Scheibe	135
5.2	Identifikation mit neuronalen Netzen	143
6	Diskussion der Ergebnisse	148
7	Zusammenfassung und Ausblick	152

1 Einleitung

Das Versagen von Bauteilen führt immer wieder zu tragischen Unfällen und zu hohen Kosten. In vielen Fällen ist ein Versagen von Bauteilen darauf zurückzuführen, daß Schädigungen, wie zum Beispiel Risse in Stahlbauteilen, nicht rechtzeitig erkannt werden. Insbesondere bei dynamischer Belastung von Bauteilen stellt ein plötzliches Versagen ohne Vorankündigung aufgrund der Schädigung des Werkstoffes ein großes Risiko dar.

Um ein plötzliches Versagen von Bauteilen zu vermeiden, sind im Flugzeugbau, Eisenbahnenwesen und Brückenbau regelmäßige Prüfungen vorgeschrieben. Bei den möglichen Untersuchungsverfahren wird zwischen zerstörenden und zerstörungsfreien Prüfmethoden unterschieden. Um hoch belastete Bauteile durch die Entnahme von Bauteilproben nicht zusätzlich zu schwächen, kommt der Entwicklung zerstörungsfreier Prüfmethoden eine immer größere Bedeutung zu. Häufig wird bei den derzeit verwendeten Prüfmethoden das dynamische System von außen angeregt und die darauf folgende Systemantwort ausgewertet. Diese Auswertung setzt die Kenntnis des Zusammenhangs der Meßwerte mit den zu identifizierenden Parametern voraus. Mathematisch führt diese Fragestellung auf ein inverses Problem. Ein inverses Problem zeichnet sich dadurch aus, daß die zu suchenden Parameter oft nur durch iterative Lösungsstrategien und Optimierungsverfahren berechnet werden können.

Im Rahmen dieser Arbeit werden verschiedene Optimierungsstrategien zur Schadenserkennung entwickelt und getestet. Die Form der Schädigung kann sowohl lochförmig als auch rißförmig sein. Die Identifikation wird auf kontaktbehaftete Risse ausgedehnt. Des Weiteren wird der Einfluß mehrerer Lastfälle und der Anzahl der Meßpunkte auf den Verlauf der Optimierung untersucht. Zusätzlich wird die Anfälligkeit der Optimierung gegenüber ungenauen Eingangsdaten getestet.

1.1 Stand der Forschung

Die zerstörungsfreie Rißerkennung ist sowohl hinsichtlich der Entwicklung geeigneter numerischer Verfahren als auch hinsichtlich effektiver experimenteller Untersuchungsmethoden an schadhaften Bauteilen ein aktuelles Forschungsgebiet. In den meisten theoretischen Arbeiten zu diesem Thema werden elastostatische Untersuchungen durchgeführt und dazugehörige Meßmethoden zum Erfassen der Verschiebungen oder der Spannung für das Systemverhalten vorgestellt. Anzumerken ist, daß auch andere physikalische Eigenschaften, wie die elektrische und

magnetische Leitfähigkeit, für die Erforschung von Materialien und Strukturen verwendet werden können.

So hat Yoshimura [48] die Verteilung des elektrischen Potentials innerhalb eines Balkens unter Anwendung der FEM berechnet und an Hand dieser eine Rißerkennung durchgeführt.

Das Messen charakteristischer Daten aus statischen Belastungsfällen wird in [1], [21], [9], [47] und [36] zur Beschreibung der inversen Problemstellung angewendet. Hwu [21] wertet in seiner Arbeit statische Dehnungen aus, um ein Loch oder einen Riß innerhalb einer Scheibe zu identifizieren, wobei acht Meßstellen der Dehnungen innerhalb der Scheibe platziert werden. Zur Beschreibung des mechanischen Verhaltens der Scheibe benutzt er ein lineares anisotropes Werkstoffgesetz. In den Arbeiten [21] und [47] wird die „Reciprocity Gap Technique“ als iterativer Lösungsalgorithmus zur Rißerkennung eingesetzt.

Rus lokalisiert in seiner Arbeit [36] eine lochförmige Schädigung in einer Scheibe, in dem er nicht nur die Verformungen der Scheibe, sondern auch deren Ableitung bezüglich des Ortes der Schädigung simuliert. Er vergleicht die Leistungsfähigkeit verschiedener Optimierungsverfahren in Bezug auf den Erfolg der Schadenserkennung miteinander.

Gallego [14] knüpft an die Arbeit von Rus [36] an, verwendet zur Erkennung von lochförmigen und rißförmigen Fehlstellen aber einen genetischen Algorithmus.

Liang [26] setzt ein neuronales Netz ein, um Schädigungen innerhalb einer quadratischen Scheibe zu identifizieren. Dazu werden mehrere Lastfälle benutzt, wobei die Meßstellen der Dehnungen innerhalb der Scheibe platziert werden. Das von Liang erstellte neuronale Netz enthält zwei verdeckte Schichten und wird nach dem Backpropagationprinzip angelernt. Untersucht werden der Einfluß unterschiedlicher Übertragungsfunktionen des Netzes und der Einfluß der Anfangswerte der Wichtungsmatrix auf den Erfolg der Schadenserkennung.

Tanaka [46] wendet das Kalman-Filter an, um Fehlstellen innerhalb elastischer Platten zu identifizieren. Dabei wird das mechanische Verhalten der Platte durch eine Randelementmethode im Frequenzbereich beschrieben.

Auch Zang ([50] und [49]) führt Simulationen im Frequenzbereich zur Schadenserkennung durch. Es werden neuronale Netze mit einem durch die „Principal Component Analysis“ reduzierten Datensatz trainiert, in [50] nach dem Backpropagationprinzip angelernt, während in [49] ein auf Kohonen [23] basierendes Netz verwendet wird. Das trainierte Netz unterscheidet zwischen einem geschädigten und einem ungeschädigten Bauteil. Die Ergebnisse werden durch experimentell gewonnene Daten gestützt.

Iterative Optimierungsverfahren zur Rißerkennung verwenden keine dynamischen Messungen, da die Simulationen mit den vorhandenen Rechenkapazitäten zu zeitaufwendig wären. Oishi [32] und Liu [27] benutzen neuronale Netze zur Erkennung von Fehlstellen, wobei Zang [50] die gewonnenen Daten aus dem Zeitbereich komprimiert, bevor er diese einem neuronalen Netz als Eingangsgrößen zur Verfügung stellt. In [49] führte Hwu eine Schadensidentifikation mit einem selbstorganisierenden Netz von Kohonen [23] durch.

Oishi [32] simuliert das zeitliche Verhalten eines durch einen Riß an der Oberfläche geschädigten Balkens mit der Finite-Elemente-Methode. Die simulierten Verschiebungen verwendet er

als Eingangsgröße für das neuronale Netz. Das trainierte Netz wird in dieser Arbeit nicht nur mit den Simulationsdaten getestet, sondern auch mit Daten, die durch Ultraschallversuche in einem Labor erzeugt worden sind.

Liu [27] identifiziert Materialparameter durch Analyse des zeitlichen Verformungsverhaltens von Faserverbundwerkstoffen. Zur Lösung des dort auftretenden Minimierungsproblems wird ein kombiniertes Verfahren, bestehend aus einem genetischen Algorithmus und einem Gradientenabstiegsverfahren, verwendet. Durch Anwendung eines neuronalen Netz führt Liu in [28] eine Schadenserkennung im Halbraum durch. Zur Simulation der auftretenden Verschiebungen an der freien Oberfläche wird ein aus der Finiten-Element-Methode und Randelementmethode bestehendes Verfahren eingesetzt.

Eine durch einen Riß geschädigte Rotorwelle untersucht He [20] mit der Finiten-Element-Methode, wobei die Schädigung in Form einer Kerbe durch einen genetischen Algorithmus identifiziert wird.

Suh [45] verwendet zum Lokalisieren einer Schädigung innerhalb einer im Frequenzbereich untersuchten Struktur eine Mischform aus neuronalem Netz und einem genetischen Algorithmus. Das neuronale Netz wird zur Klassifikation der Eigenfrequenzen eingesetzt, während der anschließend eingesetzte genetische Algorithmus den Ort der Schädigung ermittelt.

Genetische Algorithmen [24] und eine Mischform aus genetischen Algorithmen und neuronalen Netzen [45] sind auch zur Schadenserkennung in der Dynamik einsetzbar. Auf Grund der großen Datenmenge werden bei diesen iterativen Vorgehensweisen jedoch nur einfache Strukturen untersucht. Entscheidend für die Identifikation von Fehlstellen in der Dynamik ist eine brauchbare Datenkomprimierung, so daß kleine Netzarchitekturen zur Identifikation benutzt werden können.

Eine im Frequenzbereich simulierte quadratische Scheibe untersucht Stavroulakis in [43], um eine Schadenserkennung mit neuronalen Netzen durchzuführen. Die Anwendung des genetischen Algorithmus zur Schadensidentifizierung in einer Scheibe wird in [42] vorgestellt und die Simulation einer durch einen Riß geschädigten Scheibe in [40]. Dabei werden ein mögliches Schließen des Risses und auftretende Reibungsphänomene berücksichtigt. Eine Zusammenfassung dieser Arbeiten und eine gute Übersicht weiterer, älterer Veröffentlichungen bietet die Monographie [41].

1.2 Ziel und Inhalt der Arbeit

Lokale Störungen in einer Struktur führen zu einer Änderung des mechanischen Verhaltens, die als Indikator zur Fehlstellenerkennung genutzt werden kann. So können bei geschädigten Bauteilen unter gleicher Belastungssituation andere Randverformungen auftreten als bei ungeschädigten Bauteilen. Das Erkennen und nachfolgend die Lokalisierung einer Schädigung mittels gemessener charakteristischer Daten läßt sich mathematisch als ein Optimierungsproblem formulieren.

Im Rahmen dieser Arbeit werden verschiedene Optimierungsstrategien, Verfahren aus dem Ge-

biet der Soft Computing Methoden, der Filtertechnik und der klassischen Optimierung entwickelt und auf ihre Leistungsfähigkeit zur Schadenslokalisierung untersucht und verglichen.

Die Untersuchungen basieren ausschließlich auf numerischen Simulationen, für die sich der Einsatz der Randelementmethode bewährt hat. Unterschiedliche Geometrien werden unter statischer und unter dynamischer Belastung untersucht. Ein besonderer Schwerpunkt der Arbeit liegt auf der bisher noch wenig erforschten Schadenserkennung im Zeitbereich über die durch Schädigungen gestörte Wellenausbreitung im Bauteil.

Unter statischer Belastung werden zwei- und dreidimensionale Geometrien untersucht. Im Zeitbereich beschränken sich die Untersuchungen aufgrund des hohen Rechenaufwandes im Dreidimensionalen auf zweidimensionale Strukturen. Zur Fehlstellenerkennung werden hierbei nur das Kalman-Filter und neuronale Netze eingesetzt. Die Form der Schädigung innerhalb der Struktur kann sowohl lochförmig als auch rißförmig sein. Bei Identifikation rißförmiger Fehlstellen wird ein möglicher Kontakt der Rißflächen durch Gesetze der Kontaktmechanik berücksichtigt. Es wird der Einfluß mehrerer Lastfälle und der Anzahl der Meßpunkte auf den Verlauf der Optimierung untersucht. Zusätzlich wird die Anfälligkeit der Optimierung gegenüber ungenauen Eingangsdaten getestet.

2 Theoretische Grundlagen

2.1 Grundgleichungen der Elastizitätstheorie

Die elastischen Grundgleichungen der Mechanik beschreiben den Verformungs- und Spannungszustand eines elastischen Körpers. Wenn von kleinen Verformungs- und Verschiebungsgradienten ausgegangen werden kann, genügt eine lineare Betrachtung der mechanischen Eigenschaften des Körpers. Aus der Impulsbilanz ergibt sich die Gleichgewichtsbedingung in Abhängigkeit von dem Ortsvektor \mathbf{x} und der Zeit t zu

$$\sigma_{ik,k}(\mathbf{x}, t) + b_i(\mathbf{x}, t) = \rho \ddot{u}_i(\mathbf{x}, t), \quad (2.1)$$

wobei \ddot{u}_i die Beschleunigung darstellt und die Volumenkraft mit b_i bezeichnet wird. Der lineare Zusammenhang zwischen Spannungen σ_{ik} und Verzerrungen ε_{ik} wird durch das Hookesche Gesetz beschrieben

$$\sigma_{ik}(\mathbf{x}, t) = 2\mu\varepsilon_{ik}(\mathbf{x}, t) + \lambda\delta_{ik}\varepsilon_{jj}(\mathbf{x}, t), \quad (2.2)$$

wobei die Lamékoefizienten μ und λ die Materialeigenschaften darstellen, die über

$$\mu = G = \frac{E}{2(1+\nu)}, \quad (2.3)$$

$$\lambda = \frac{\nu E}{(1-2\nu)(1+\nu)} \quad (2.4)$$

mit der Poissonschen Zahl ν und dem Elastizitätsmodul E verknüpft sind.

Zwischen Verschiebungen u_{ik} und den Verzerrungen ε_{ik} gilt bei Annahme kleiner Verschiebungsgradienten der lineare Zusammenhang

$$\varepsilon_{ik}(\mathbf{x}, t) = \frac{1}{2}(u_{i,k}(\mathbf{x}, t) + u_{k,i}(\mathbf{x}, t)). \quad (2.5)$$

Die drei Grundgleichungen (2.1), (2.2) und (2.5) können zu den Navierschen Gleichungen zusammengefaßt werden

$$\mu u_{i,jj}(\mathbf{x}, t) + (\mu + \lambda) u_{j,ij}(\mathbf{x}, t) + b_i(\mathbf{x}, t) = \rho \ddot{u}_i(\mathbf{x}, t). \quad (2.6)$$

Werden nur statische Problemstellungen behandelt, verlieren die einzelnen Komponenten ihre Zeitabhängigkeit und die Trägheitsterme $\rho \ddot{u}_i(\mathbf{x}, t)$ werden zu null.

2.2 Zeitabhängige Randintegralgleichung

Eine Möglichkeit, Verfahren zur Lösung von partiellen Differentialgleichungen zu finden, ist die Methode der gewichteten Residuen. Dabei werden die Differentialgleichungen aus ihrer starken Form in eine globale Integralaussage überführt. Die Differentialgleichung wird mit einer Wichtungsfunktion multipliziert, integriert und dieses gewichtete Integral zu Null gesetzt. Der Gedanke hinter dieser Methode ist, die Differentialgleichungen nicht exakt zu erfüllen, sondern nur zu fordern, daß der gewichtete Fehler in dem zu integrierenden Gebiet im Mittel verschwindet.

Werden alle auftretenden Ableitungen durch partielles Integrieren auf die bekannte Wichtungsfunktion „verschoben“, ergeben sich je nach Wahl der Wichtungsfunktion verschiedene numerische Verfahren, u.a. die Finite-Element Methode oder die Randelementmethode.

2.2.1 Fundamentallösung

Zur Herleitung der Randintegralgleichungen, die die Basis der Randelementmethode bilden, wird als Wichtungsfunktion die Fundamentallösung der Differentialgleichung verwendet. Diese ist laut [2] und [37] dadurch charakterisiert, daß sie in allen Punkten \mathbf{x} die Lösung der homogenen Gleichung darstellt, mit Ausnahme des Quellpunktes $\boldsymbol{\xi}$, in dem sie eine Singularität aufweist. Physikalisch können die Fundamentallösungen als Antworten eines unendlichen Kontinuums auf punktförmige Einheitsimpulsanregungen interpretiert werden. Mathematisch wird diese Anregung durch

$$\mathbf{b}^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) = \delta(t - \tau) \delta(\mathbf{x} - \boldsymbol{\xi}) \mathbf{e}^{(j)} \quad (2.7)$$

oder mit dem Kroneckersymbol δ_{ij} in Indexschreibweise

$$b_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) = \delta(t - \tau) \delta(\mathbf{x} - \boldsymbol{\xi}) \delta_{ij} \quad (2.8)$$

beschrieben. Die Fundamentallösung ist die Systemantwort in dem Feldpunkt \mathbf{x} zur Zeit t infolge eines im Quellpunkt $\boldsymbol{\xi}$ zur Zeit τ angreifenden Dirac-Impulses, der in Richtung des Einheitsvektors $\mathbf{e}^{(j)}$ zeigt.

Das Kontinuum reagiert auf die Anregungen $\mathbf{b}^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)$ mit den Verschiebungen $u_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)$. Da eine Impulsanregung in eine Richtung (j) Verschiebungen in alle Raumrichtungen zur Folge hat, besitzt $u_i^{(j)}$ die Form eines Tensors.

Die Fundamentallösung $u_i^{(j)}$ für ein elastisches, unendlich ausgedehntes Kontinuum ist die Lösung der speziellen Navierschen Gleichung (2.6)

$$\rho \ddot{u}_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) - \mu u_{i,kk}^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) - (\mu + \lambda) u_{k,ik}^{(j)}(\mathbf{x}, t) = \delta(t - \tau) \delta(\mathbf{x} - \boldsymbol{\xi}) \delta_{ij} . \quad (2.9)$$

Mit Cauchys Theorem (2.13), dem Hookeschen Gesetz (2.2) und der Verschiebungskinetik (2.5) kann, ausgehend von der Verschiebungsfundamentallösung $u_i^{(j)}$, (2.9) der zugehörige Spannungsvektor $t_i^{(j)}$ ermittelt werden. Die Herleitung der elastodynamischen Fundamentallösungen für den dreidimensionalen Vollraum ist in [37] beschrieben. Für den zweidimensionalen Fall sind die fundamentalen Lösungen in [2] angegeben.

Die Fundamentallösungen zeitabhängiger Probleme besitzen zusätzlich zu den oben angeführten, noch die drei wesentlichen Eigenschaften der Translation, der Reziprozität und der Kausalität.

Die Translation besagt, daß lediglich die Orts- bzw. Zeitdifferenz zwischen Beobachtungs- und Belastungspunkt entscheidend ist. Durch die Reziprozität sind Quell- und Feldpunkt sowie Beobachtungs- und Belastungszeitpunkt vertauschbar, und die Kausalität stellt sicher, daß an einem Beobachtungspunkt vor Erreichen der Kompressionswelle keine Reaktion hervorgerufen wird.

2.2.2 Herleitung mit der Methode der gewichteten Residuen

Das gewichtete Residuum der Navierschen Gleichungen (2.6) mit den Fundamentallösung $u_i^{(k)}$ als Wichtungsfunktion hat die Form

$$\int_0^t \int_{\Omega} [\mu u_{i,jj}(\mathbf{x}, \tau) + (\mu + \lambda) u_{j,ij}(\mathbf{x}, \tau) + b_i(\mathbf{x}, \tau) - \rho \ddot{u}_i] u_i^{(k)} d\Omega_x d\tau = 0 \quad (2.10)$$

Werden durch partielle Integrationen alle partiellen Ableitungen auf die bekannte Wichtungsfunktion $u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)$ verschoben, so ergibt sich

$$\begin{aligned} & \int_0^t \oint_{\Gamma} [t_i(\mathbf{x}, t) u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t) - u_i(\mathbf{x}, \tau) t_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)] d\Gamma_x \\ & + \int_{\Omega} [b_i(\mathbf{x}, \tau) u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)] d\Omega_x \\ & + \int_{\Omega} [\mu u_{i,jj}^{(k)} + (\mu + \lambda) u_{j,ij}^{(k)} - \rho \ddot{u}_i^{(k)}] u_i(\mathbf{x}, \tau) d\Omega_x \} d\tau \\ & - \int_{\Omega} [\dot{u}_i(\mathbf{x}, \tau) u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) - u_i(\mathbf{x}, \tau) \frac{\partial u_i^{(k)}}{\partial \tau}]_{\tau=0}^t d\Omega_x = 0. \end{aligned} \quad (2.11)$$

Betrachtet man die Definition (2.9) der Fundamentalgleichung, d.h. die Filtereigenschaft der Dirac Funktion und die Kausalität der Fundamentallösung, d.h. für jeden Punkt $\mathbf{x} \neq \boldsymbol{\xi}$

$u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) = 0$ und $\frac{\partial u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)}{\partial \tau} = 0$, ergibt sich für alle Punkte $\boldsymbol{\xi} \in \Omega$

$$\begin{aligned} u_i(\boldsymbol{\xi}) = & \int_0^t \oint_{\Gamma} [t_i(\mathbf{x}, \tau) u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) - u_i(\mathbf{x}, \tau) t_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)] d\Gamma \\ & + \int_{\Omega} [b_i(\mathbf{x}, \tau) u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)] d\Omega \} d\tau \\ & - \int_{\Omega} [\dot{u}_i(\mathbf{x}, 0) u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, 0) - u_i(\mathbf{x}, 0) \frac{\partial}{\partial \tau} u_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)]_{\tau=0} d\Omega. \end{aligned} \quad (2.12)$$

Dabei sind durch

$$t_i(\mathbf{x}, \tau) = \sigma_{ij}(\mathbf{x}, \tau) n_j(\mathbf{x}), \quad (2.13)$$

$$t_i^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) = \sigma_{ij}^{(k)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) n_j(\mathbf{x}) \quad (2.14)$$

$$(2.15)$$

die Randspannungsvektoren der gesuchten Lösung bzw. der Fundamentallösung definiert.

2.2.3 Randintegralgleichung

Für die Berechnung der Verschiebungen $u_j(\boldsymbol{\xi}, t)$ muß die vollständige Kenntnis der Randwerte der Spannungen und Verschiebungen vorausgesetzt werden. Bei einem korrekt gestellten Problem muß immer jeweils eine Randgröße bekannt sein.

Zur Berechnung der unbekannten Randwerte wird der Quellpunkt $\boldsymbol{\xi}$ auf den Rand verschoben. Fallen dabei $\boldsymbol{\xi}$ und der Integrationspunkt \mathbf{x} auf dem Rand zusammen, wird der Abstand $r = |\mathbf{x} - \boldsymbol{\xi}|$ zu $r = 0$ und somit werden die Fundamentallösungen singulär. und zwar Die Fundamentallösung von $u_i^{(j)}$ wird schwach singulär und die für $t_i^{(j)}$ stark singulär. Schwach singulär bedeutet für zweidimensionale Problemstellungen eine Singularität wie $\ln(r)$, stark singulär eine wie $\frac{1}{r}$.

Zur Bestimmung des Verhaltens der Integralgleichung (2.12) am singulären Punkt $\mathbf{x} = \boldsymbol{\xi}$ ist bei der Integration eine Grenzwertbetrachtung notwendig. Die Umgebung dieses Punktes $\boldsymbol{\xi}$ wird um Ω_ϵ , für den zweidimensionalen Fall mit dem kreisförmigen Rand Γ_ϵ , erweitert, so daß $\boldsymbol{\xi}$ weiterhin im betrachteten Gebiet Ω liegt (siehe Abbildung 2.1). Verkleinert man nun den Radius ϵ von Γ_ϵ , gleichzusetzen mit dem Grenzübergang $\epsilon \rightarrow 0$, wird Gleichung (2.12) zu

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \int_{\Gamma - \Gamma_\epsilon} (t_i * u_i^{(j)} - t_i^{(j)} * u_i) d\Gamma + \lim_{\epsilon \rightarrow 0} \int_{\Gamma_\epsilon} (t_i * u_i^{(j)} - t_i^{(j)} * u_i) d\Gamma \\ + \int_{\Omega} (b_i * u_i^{(j)} + \varrho \{ v_{i0} u_i^{(j)} + u_{i0} \dot{u}_i^{(j)} \}) d\Omega = u_j(\boldsymbol{\xi}, t). \end{aligned} \quad (2.16)$$

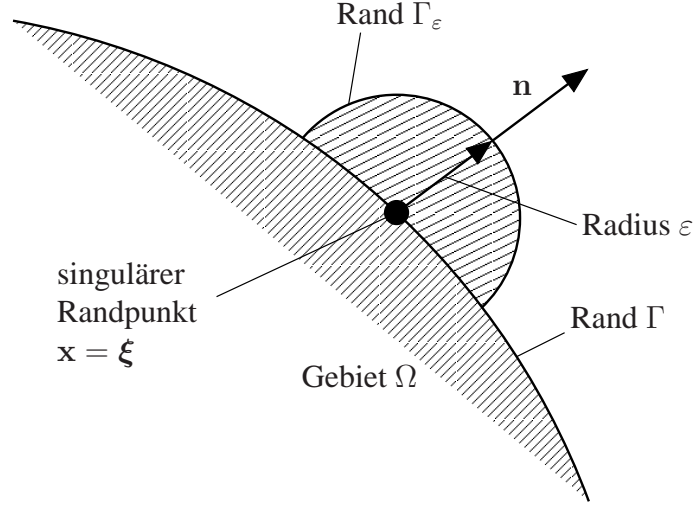


Abbildung 2.1: Erweiterter Rand zur Integration

In Gleichung (2.12) kennzeichnet der Operator $*$ eine Faltung hinsichtlich der Zeit. Diese ist durch

$$t_i * u_i = \begin{cases} 0 & t < 0 \\ \int_0^t t_i(t - \tau) u_i(\tau) d\tau & t \geq 0 \end{cases} \quad (2.17)$$

definiert. Die Terme des ersten Integrals aus Gleichung (2.16), die die schwach singulären Kerne $u_i^{(j)}$ beinhalten, existieren als uneigentliche Integrale, d.h. sie besitzen trotz ihrer Singularität einen endlichen Grenzwert. Die stark singulären Anteile lassen sich im Sinne des Cauchyschen Hauptwertes bestimmen. Beide Anteile lassen sich mit entsprechenden numerischen Integrationsverfahren bestimmen.

Aus dem zweiten Summanden der Gleichung (2.16) ergibt sich der so genannte integralfreie Term. Zusammen mit der rechten Seite $u_j(\xi, t)$ entsteht die von der Zeit unabhängige Randmatrix $c_i^{(j)}$. Aufgrund seiner schwachen Singularität verschwindet die erste Faltung im Integral im Grenzübergang $\varepsilon \rightarrow 0$. Die Randmatrix lautet damit nach [2] im zweidimensionalen Raum und mit dem Außenwinkel der Berandung $(\varphi_2 - \varphi_1)$

$$c_i^{(j)}(\xi) = \int_{\Gamma_\varepsilon} t_i^{(j)}(\mathbf{x}, \xi) d\Gamma = -\frac{1}{2\pi} \cdot \left\{ (\varphi_2 - \varphi_1) \delta_{ij} + \frac{\mu + \lambda}{2\mu + \lambda} \cdot \frac{1}{2} \cdot \begin{bmatrix} \sin 2\varphi_2 - \sin 2\varphi_1 & -\cos 2\varphi_2 + \cos 2\varphi_1 \\ -\cos 2\varphi_2 + \cos 2\varphi_1 & -\sin 2\varphi_2 + \sin 2\varphi_1 \end{bmatrix} \right\}. \quad (2.18)$$

Für die Gleichung (2.12) ergibt sich nun

$$\begin{aligned} c_i^{(j)}(\boldsymbol{\xi})u_i(\boldsymbol{\xi}, t) &= \int_{\Gamma} t_i(\mathbf{x}, \tau) * u_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) d\Gamma - \oint_{\Gamma} t_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) * u_i(\mathbf{x}, \tau) d\Gamma \\ &+ \int_{\Omega} \left(b_i(\mathbf{x}, \tau) * u_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) + \varrho \left\{ v_{i0}(\mathbf{x})u_i^{(j)}(\mathbf{x}, t) + u_{i0}(\mathbf{x})\dot{u}_i^{(j)}(\mathbf{x}, t) \right\} \right) d\Omega . \end{aligned} \quad (2.19)$$

Das C auf dem zweiten Integral kennzeichnet den Cauchyschen Hauptwert. Bei einer Problemstellung ohne Volumenkräfte und mit verschwindenden Anfangsbedingungen, vereinfacht sich Gleichung (2.19) zu

$$c_i^{(j)}(\boldsymbol{\xi})u_i(\boldsymbol{\xi}, t) = \int_{\Gamma} t_i(\mathbf{x}, \tau) * u_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) d\Gamma - \oint_{\Gamma} t_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) * u_i(\mathbf{x}, \tau) d\Gamma . \quad (2.20)$$

Diese Randintegralgleichung ist die Ausgangsgleichung der Randelementmethode.

2.2.4 Hypersinguläre Randintegralgleichung

In Abbildung 2.2 ist die Diskretisierung eines Risses dargestellt. Die obere Rißseite wird mit Γ^+ , die untere mit Γ^- und der äußere Rand der Scheibe mit Γ^0 bezeichnet. Die zuvor betrachtete Herleitung der Randintegralgleichung (2.20) gilt grundsätzlich für alle Ränder eines Gebietes, d.h. auch für innere Berandungen. Bei der Randintegralformulierung eines schmalen Risses treten allerdings, wie in [13] beschrieben, Probleme auf, da die beiden Rißseiten und damit auch die Diskretisierungspunkte fast übereinander liegen. Bei Auswertung der Integralgleichung (2.20) ergeben sich bei Betrachtung der Rißoberseite Γ^+ und der Rißunterseite Γ^- identische Einträge in den Einflußmatrizen, die zu einem singulären Gleichungssystem führen. Um dies zu vermeiden, wird für einen Rißrand die zu beschreibende Verschiebungsintegralgleichung durch die Spannungsintegralgleichung ersetzt. Somit stehen für einen Rißpunkt zwei voneinander unabhängige Gleichungen zu Verfügung.

Aus dem Cauchy Theorem, dem Hookeschen Gesetz (2.2) und der Verschiebungsskinematik (2.5) folgt die funktionale Beziehung zwischen Spannungs- und Verschiebungsvektor zu

$$t_i = \left(\lambda \delta_{ik} u_{j,j} + \mu (u_{i,k} + u_{k,i}) \right) n_k . \quad (2.21)$$

Ausgehend von der Verschiebungsintegralgleichung eines Punktes innerhalb des Gebietes Ω

$$u_i(\boldsymbol{\xi}, t) = \int_{\Gamma} t_i(\mathbf{x}, \tau) * u_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) d\Gamma - \int_{\Gamma} t_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) * u_i(\mathbf{x}, \tau) d\Gamma , \quad (2.22)$$

bei der $\Gamma = \Gamma^0 \cup \Gamma^+ \cup \Gamma^-$ gilt, kann durch Differenzieren dieser Integralgleichung nach $\boldsymbol{\xi}_{,i}$ und $\boldsymbol{\xi}_{,j}$ und Einsetzen in Gleichung (2.21) die Spannungsintegralgleichung für Innenpunkte $\boldsymbol{\xi} \in \Omega$

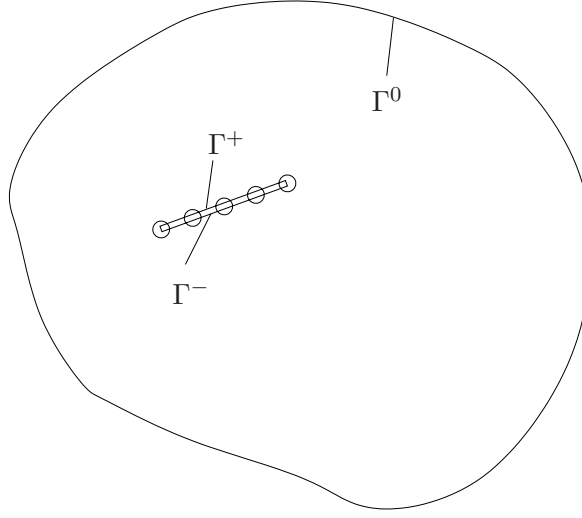


Abbildung 2.2: Rißberandung mit übereinander liegenden Diskretisierungspunkten

formuliert werden.

$$t_j(\boldsymbol{\xi}, t) = \int_{\Gamma} [t_i(\mathbf{x}, t) * d_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau)] d\Gamma - \int_{\Gamma} [s_i^{(j)}(\mathbf{x}, t) * u_i(\mathbf{x}, \boldsymbol{\xi}, t, \tau)] d\Gamma . \quad (2.23)$$

Anschließend wird $\boldsymbol{\xi}$ auf den Rißrand verschoben, wodurch die Integralkerne der Gleichung (2.23) stark singular bzw. hypersingular werden. Die numerische Umsetzung der hypersingulären und stark singulären Integrale können in [17] nachgelesen werden.

$$\frac{1}{2} t_j(\boldsymbol{\xi}, t) = \int_{\Gamma} [d_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) * t_j(\mathbf{x}, t)] d\Gamma - \int_{\Gamma} [s_i^{(j)}(\mathbf{x}, \boldsymbol{\xi}, t, \tau) * u_j(\mathbf{x}, t)] d\Gamma \quad (2.24)$$

Solange der Riß geschlossen bleibt, liegen die Elemente der Rißflanken übereinander, so daß bei Auswertung der Integralgleichung ein weiterer integralfreier Term entsteht [33]. Bei Öffnung des Risses verschwindet dieser jedoch.

Damit steht am Riß eine zweite von (2.20) unabhängige Integralgleichung zur Verfügung. Zur Lösung des Gesamtsystems wird für den Außenrand und für den oberen Rand des Risses Γ^+ die Verschiebungsgleichung (2.20) benutzt. Bei Lage des Kollokationspunktes auf dem unteren Rißrand Γ^- wird jedoch die Spannungsgleichung (2.22) ausgewertet. Somit kann ein singuläres Gleichungssystem zur Lösung des Gesamtsystems umgangen werden.

2.2.5 Randelementformulierung

In Abschnitt 2.2.3 wurden Randintegralgleichungen für elastodynamische Probleme hergeleitet. Für diese Gleichungen (2.20) und (2.24) gibt es im Allgemeinen keine analytische Lösung. Es ist jedoch möglich, sie mit Diskretisierungsverfahren, d.h. mit der Randelementmethode, zu lösen. Diese impliziert nach [2] folgende Vorgehensweise:

Zunächst wird der Untersuchungszeitraum und der Rand des Gebietes diskretisiert. Es folgt eine Approximation der Geometrie des Randes und der Randzustandsgrößen durch Ansatzfunktionen in Ort und Zeit. Die Randintegrale werden daraufhin bezüglich Ort und Zeit ausgewertet, wobei die Integration für den Ort numerisch, hinsichtlich der Zeit jedoch analytisch erfolgt [37].

Nach der Randintegration wird ein lineares Gleichungssystem zur Berechnung der unbekannten Randwerte aufgestellt. Dies geschieht durch Punktkollokation, d.h. durch die punktweise Erfüllung der Randintegralgleichung in den Knoten der Ansatzfunktionen. Abschließend wird das so entstandene algebraische Gleichungssystem für den aktuellen Zeitschritt gelöst.

Dieses Vorgehen führt bei Berücksichtigung der Translationseigenschaft der Matrix, wie in [37] gezeigt, auf eine Rekursionsformel für die diskrete elastische Randintegralgleichung im Zeitbereich

$$\mathbf{C}^1 \mathbf{y}^n = \mathbf{D}^1 \bar{\mathbf{y}}^n + \mathbf{R}^n. \quad (2.25)$$

Hierfür ist eine äquidistante Zeitschrittweite δt notwendig. In dieser Gleichung stehen in dem Vektor \mathbf{y}^n die unbekannten und in $\bar{\mathbf{y}}^n$ die bekannten Randwerte, also bei der Approximation der Gleichung (2.20) die Verschiebungen \mathbf{u} und Spannungen \mathbf{t} auf dem jeweiligen Teil des Randes Γ für den aktuellen Zeitschritt n

$$\mathbf{y}^n = \begin{bmatrix} \mathbf{t}_{\Gamma_u} \\ \mathbf{u}_{\Gamma_t} \end{bmatrix}^n \quad \text{und} \quad \bar{\mathbf{y}}^n = \begin{bmatrix} \bar{\mathbf{u}}_{\Gamma_u} \\ \bar{\mathbf{t}}_{\Gamma_t} \end{bmatrix}^n. \quad (2.26)$$

In den Matrizen \mathbf{C}^1 und \mathbf{D}^1 aus Gleichung (2.25) sind die Spannungs- und Verschiebungsfundamentallösungen des ersten Zeitschrittes zusammengefaßt, wobei die Matrizen \mathbf{U} und $\check{\mathbf{T}}$ die Faltung der Fundamentallösungen mit den Ansatzfunktionen für die Spannungs- und Verschiebungsvektoren enthalten

$$\mathbf{C}^1 = [-\mathbf{U}_{\Gamma_u} | \check{\mathbf{T}}_{\Gamma_t}]^1 \quad \text{und} \quad \mathbf{D}^1 = [-\check{\mathbf{T}}_{\Gamma_u} | \mathbf{U}_{\Gamma_t}]^1. \quad (2.27)$$

Durch die Matrix \mathbf{R}^n in Gleichung (2.25) wird der Einfluß der Lastgeschichte bis zu dem Zeitschritt n beschrieben:

$$\mathbf{R}^n = \sum_{k=2}^n \left(\mathbf{U}^k \mathbf{t}^{n-k+1} - \mathbf{T}^k \mathbf{u}^{n-k+1} \right). \quad (2.28)$$

Die Addition in Gleichung (2.28) wird bis zu dem Zeitschritt n durchgeführt, der bei einer zweidimensionalen Problemstellung dem Ende des bisher betrachteten Zeitintervalls entspricht.

In Gleichung (2.25) sind also lediglich die gesuchten Randwerte \mathbf{y}^n unbekannt. Durch Multiplikation dieser Gleichung mit der Inversen $(\mathbf{C}^1)^{-1}$ können die unbekannten Größen \mathbf{y}^n berechnet werden.

2.3 Kontaktmechanik

Die Kontaktmechanik betrachtet das Problem des Kontaktes zweier elastischer Körper. Als Signorinikontaktbedingung wird die mathematische Darstellung des unilateralen Kontaktes bezeichnet. Auch das Schließen eines Risses kann mittels der Kontaktmechanik beschrieben werden. Bei Auftreten eines Kontaktes entsteht eine zur Kontaktfläche senkrecht wirkende Normalspannung t_n , wobei die Verschiebung in Normalenrichtung u_n null ist. Bei Vernachlässigung der Reibung entfällt die Tangentialspannung an der Kontaktfläche. Die Verschiebung in diese Richtung ist möglich. Bei Aufhebung des Kontaktes verschwindet die Normalspannung senkrecht zur Kontaktfläche unter Auftreten einer relativen Normalverschiebung Δu_n als Differenz der Verschiebungen zweier gegenüberliegender Randpunkte.

2.3.1 Formulierung des Kontaktes

Die Gesetzmäßigkeiten des Kontaktes lassen sich durch folgende Komplementaritätsbedingung (LCP) beschreiben

$$\Delta u_n \geq 0, \quad (2.29)$$

$$t_n \geq 0, \quad (2.30)$$

$$\Delta u_n t_n = 0. \quad (2.31)$$

Gleichung (2.31) besagt, daß immer eine der beiden Größen Normalverschiebung oder Normalspannung den Wert null hat. Die Randelementformulierung läßt sich in folgender Form als Ausgangspunkt zum Einbau der Kontaktbedingungen schreiben:

$$\begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_a \\ \Delta \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y}_a \\ \mathbf{t}_c \end{bmatrix}, \quad (2.32)$$

wobei in \mathbf{t}_c die auftretenden Spannungen an der Kontaktfläche enthalten sind. Zur Überführung dieser Randelementformulierung zu einem linearen Komplementaritätsproblem werden alle bekannten Randgrößen \mathbf{x}_a und unbekannten Randgrößen \mathbf{y}_a des äußeren Randes, als auch die unbekannten relativen Verschiebungen des Risses $\Delta \mathbf{u}$ in Tangentenrichtung des Risses aus Gleichung (2.32) eliminiert. Dazu wird die erste Gleichung aus (2.32) nach \mathbf{x}_a aufgelöst

$$\mathbf{x}_a = \mathbf{T}_{11}^{-1}(-\mathbf{T}_{12}\Delta \mathbf{u} + \mathbf{U}_{11}\mathbf{y}_a + \mathbf{U}_{12}\mathbf{t}_c) \quad (2.33)$$

und in den zweiten Teil der Gleichung (2.32) eingesetzt. Es ergibt sich folgender Ausdruck

$$\mathbf{T}_{21}(\mathbf{T}_{11}^{-1}(-\mathbf{T}_{12}\Delta \mathbf{u} + \mathbf{U}_{11}\mathbf{y}_a + \mathbf{U}_{12}\mathbf{t}_c)) + \mathbf{T}_{22}\Delta \mathbf{u} = \mathbf{U}_{21}\mathbf{y}_a + \mathbf{U}_{22}\mathbf{t}_c \quad (2.34)$$

Zur Vereinfachung werden folgende FormelAusdrücke zusammengefaßt:

$$\bar{\mathbf{T}} = \mathbf{T}_{21}\mathbf{T}_{11}^{-1}(-\mathbf{T}_{12}) + \mathbf{T}_{22}, \quad (2.35)$$

$$\bar{\mathbf{U}} = \mathbf{T}_{21}\mathbf{T}_{11}^{-1}\mathbf{U}_{12} - \mathbf{U}_{22}, \quad (2.36)$$

$$\bar{\mathbf{q}} = (-\mathbf{T}_{21}\mathbf{T}_{11}^{-1}\mathbf{U}_{11} + \mathbf{U}_{21})\mathbf{y}_a \quad (2.37)$$

Gleichung (2.32) kann nun als

$$\bar{\mathbf{T}}\Delta\mathbf{u} + \bar{\mathbf{U}}\mathbf{t}_c = \bar{\mathbf{q}} \quad (2.38)$$

dargestellt werden. In diesem Ausdruck beziehen sich \mathbf{t}_c und $\Delta\mathbf{u}$ auf das globale Koordinatensystem. Es wird nun eine Koordinatentransformation mit der Transformationsmatrix \mathbf{C} durchgeführt, die Gleichung (2.38) in eine Darstellung im lokalen Rißkoordinatensystem darstellt.

$$\hat{\mathbf{T}}\Delta\hat{\mathbf{u}}_a + \hat{\mathbf{U}}\hat{\mathbf{t}}_c = \bar{\mathbf{q}} \quad (2.39)$$

Für die transformierten Größen gelten folgende Abkürzungen und Zusammenhänge

$$\begin{aligned} \Delta\hat{\mathbf{u}}_a &= \mathbf{C}\Delta\mathbf{u}, \\ \hat{\mathbf{t}}_c &= \mathbf{C}\mathbf{t}_c, \\ \hat{\mathbf{T}} &= \bar{\mathbf{T}}\mathbf{C}^T, \\ \hat{\mathbf{U}} &= \bar{\mathbf{U}}\mathbf{C}^T. \end{aligned} \quad (2.40)$$

Nach Gleichung (2.39) ist nur der Teil der Gleichung ein Komplementaritätspaar, der die Größen senkrecht zu der Rißfläche beschreibt. In einem weiteren Eliminationsschritt wird Gleichung (2.39) auf diese Paare reduziert. Dazu wird sie folgendermaßen umgeschrieben

$$\begin{pmatrix} \tilde{\mathbf{T}}_{11} & \tilde{\mathbf{T}}_{12} \\ \tilde{\mathbf{T}}_{21} & \tilde{\mathbf{T}}_{22} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{u}_{c1} \\ \Delta\mathbf{u}_{c2} \end{pmatrix} + \begin{pmatrix} \tilde{\mathbf{U}}_{11} & \tilde{\mathbf{U}}_{12} \\ \tilde{\mathbf{U}}_{21} & \tilde{\mathbf{U}}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{t}_{c1} \\ \mathbf{t}_{c2} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{q}}_1 \\ \bar{\mathbf{q}}_2 \end{pmatrix}, \quad (2.41)$$

in der $\Delta\mathbf{u}_{c1}$ die Verschiebungen des Risses in Tangentenrichtung und $\Delta\mathbf{u}_{c2}$ die Verschiebungen des Risses in Normalenrichtung beschreiben. Die Spannungen des Risses werden mit \mathbf{t}_{c1} in Tangentenrichtung angegeben und mit \mathbf{t}_{c2} in Normalenrichtung. Die erste Gleichung von (2.41) kann nach \mathbf{u}_{c1} aufgelöst

$$\Delta\mathbf{u}_{c1} = \tilde{\mathbf{T}}_{11}^{-1} \left(-\tilde{\mathbf{T}}_{12}\Delta\mathbf{u}_{c2} - \tilde{\mathbf{U}}_{11}\mathbf{t}_{c1} - \tilde{\mathbf{U}}_{12}\mathbf{t}_{c2} + \bar{\mathbf{q}}_1 \right). \quad (2.42)$$

und in den zweiten Teil eingesetzt werden. Mit den Abkürzungen

$$\mathbf{H} = \tilde{\mathbf{T}}_{21}\tilde{\mathbf{T}}_{11}^{-1} \left(-\tilde{\mathbf{T}}_{12} \right) + \tilde{\mathbf{T}}_{22}, \quad (2.43)$$

$$\mathbf{G} = \tilde{\mathbf{T}}_{21}\tilde{\mathbf{T}}_{11}^{-1} \left(-\tilde{\mathbf{U}}_{12} \right) + \tilde{\mathbf{U}}_{22}, \quad (2.44)$$

$$\mathbf{g} = -\tilde{\mathbf{T}}_{21}\tilde{\mathbf{T}}_{11}^{-1} \left(-\tilde{\mathbf{U}}_{11}\mathbf{t}_{c1} + \bar{\mathbf{q}}_1 \right) - \tilde{\mathbf{U}}_{21}\mathbf{t}_{c1} + \bar{\mathbf{q}}_2 \quad (2.45)$$

kann Gleichung (2.41) vereinfacht werden zu

$$\mathbf{H}\Delta\mathbf{u}_{c2} + \mathbf{G}\mathbf{t}_{c2} = \mathbf{g}. \quad (2.46)$$

Da in den Untersuchungen die auftretende Reibung nicht weiter berücksichtigt wird, kann für Gleichung (2.45) geschrieben werden:

$$\mathbf{g} = -\tilde{\mathbf{T}}_{21}\tilde{\mathbf{T}}_{11}^{-1}\bar{\mathbf{q}}_1 + \bar{\mathbf{q}}_2. \quad (2.47)$$

Wird Gleichung (2.46) linksseitig mit \mathbf{H}^{-1} multipliziert, so ergibt sich das folgende lineare Komplementaritätsproblem (LCP)

$$\begin{aligned}\Delta \mathbf{u}_{c2} - \mathbf{M} \mathbf{t}_{c2} &= \mathbf{q}, \\ \Delta \mathbf{u}_{c2} &\geq \mathbf{0}, \\ \mathbf{t}_{c2} &\geq \mathbf{0}, \\ \Delta \mathbf{u}_{c2}^T \mathbf{t}_{c2} &= 0,\end{aligned}\tag{2.48}$$

$$\mathbf{M} = -\mathbf{H}^{-1} \mathbf{G}\tag{2.49}$$

wird dabei als Flexibilitätsmatrix bezeichnet und \mathbf{q} steht für den Ausdruck

$$\mathbf{q} = \mathbf{H}^{-1} \mathbf{g}.\tag{2.50}$$

Wenn das LCP der Gleichung (2.48) gelöst worden ist, können im Anschluss die restlichen Unbekannten ermittelt werden.

2.3.2 Lemke-Verfahren

Die Lösung eines linearen Komplementaritätsproblems (LCP) wie in (2.48) aufgestellt, kann iterativ mit dem Lemke-Algorithmus erfolgen. Einzelheiten und weitere Anwendungen des Lemke-Verfahrens können in [31] nachgelesen werden. Ausgangspunkt ist das zu lösende LCP nach (2.51).

$$\begin{aligned}\mathbf{w} - \mathbf{M} \mathbf{z} &= \mathbf{q}, \\ \mathbf{w} &\geq \mathbf{0}, \\ \mathbf{z} &\geq \mathbf{0}, \\ \mathbf{w}^T \mathbf{z} &= 0,\end{aligned}\tag{2.51}$$

Gesucht ist die Lösung $\mathbf{x} = (\mathbf{w}^T, \mathbf{z})$, die das LCP (2.51) erfüllen. Hierzu wird das zu beschreibende Gleichungssystem erweitert zu

$$\begin{aligned}\mathbf{w} - \mathbf{M} \mathbf{z} - \mathbf{h} z_0 &= \mathbf{q}, \\ \mathbf{w} &\geq \mathbf{0}, \\ \mathbf{z} &\geq \mathbf{0}, \\ z_0 &\geq 0, \\ \mathbf{w}^T \mathbf{z} &= 0,\end{aligned}\tag{2.52}$$

mit $\mathbf{h} = (1, 1, \dots, 1)^T$ und der zusätzlichen Variablen z_0 , die im Zusammenhang mit dem Lemke-Algorithmus als zusätzliche Basisvariable bezeichnet wird. z_0 kann als fiktive Vorspannung oder Versetzung gesehen werden, die im Laufe der Iterationen schrittweise herausgenommen werden muß. Unter der Vorgabe, daß $z_0 = -\min_i(q_i)$ gelten soll, ist $\mathbf{w} = \mathbf{q} + \mathbf{h} z_0$ und $\mathbf{z} = \mathbf{0}$ eine Lösung des LCP.

Basis	z_1	\cdot	z_l	\cdot	z_m	z_0	q
w_1	$-M_{11}$	\cdot	$-M_{1l}$	\cdot	$-M_{1m}$	-1	q_1
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
w_s	$-M_{s1}$	\cdot	$-M_{sl}$	\cdot	$-M_{sm}$	-1	q_s
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
w_m	$-M_{m1}$	\cdot	$-M_{ml}$	\cdot	$-M_{mm}$	-1	q_m

Tabelle 2.1: Lineares Komplementaritätsproblem in Tableauform

Das LCP kann in Tabelle 2.1 in Tableauform dargestellt werden. Dabei sind M_{ij} ($i, j = 1, \dots, m$) die Komponenten der Matrix M . q_i ($i = 1, \dots, m$) sind die Einträge des Vektors q . Als Basisvariablen werden w_i ($i = 1, \dots, m$) und als Nichtbasisvariablen werden z_i ($i = 1, \dots, m$) bezeichnet.

Ziel ist es, durch Tauschen der Basis- und Nichtbasisvariablen eine Lösung des LCP zu erzeugen. Ist $q_i > 0 \forall i = 1, \dots, m$, so ist $w = q$ eine Basislösung des LCP und der Algorithmus kann beendet werden. Gilt für einen oder mehrere Einträge $q_i < 0$, wird das Verfahren fortgeführt. Das Tauschen der Basis- und der Nichtbasisvariablen erfolgt analog zum Simplexalgorithmus über Pivolisieren.

Der erste Pivotschritt ist fest vorgegeben. Die Pivotspalte ist die Spalte, in der z_0 steht. z_0 wird zur Basisvariablen. Die Pivotzeile wird gemäß der Vorschrift $q_s = \min_i \{q_s | q_s < 0\}$ bestimmt. Die Variable w_s wird zur Nichtbasisvariablen. In Tabelle 2.1 ist das Pivotelement -1 eingrahmt. Nach Durchführung des ersten Pivotschrittes mit den mathematischen Umformungen

$$q'_s = -q_s, \quad (2.53)$$

$$q'_i = q_i - q_s M_{ik}, \quad i \neq s, \quad (2.54)$$

$$M'_{sj} = -M_{sj}, \quad (2.55)$$

$$M'_{ij} = M_{ij} - M_{sj} M_{ik}, \quad i \neq s, \quad (2.56)$$

entsteht das Tableau in Tabelle 2.2.

Basis	z_1	\cdot	z_k	\cdot	z_m	z_0	q
w_1	$-M'_{11}$	\cdot	$-M'_{1k}$	\cdot	$-M'_{1m}$	0	q'_1
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
z_0	$-M'_{s1}$	\cdot	$-M'_{sk}$	\cdot	$-M'_{sm}$	1	q'_s
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
w_m	$-M'_{m1}$	\cdot	$-M'_{mk}$	\cdot	$-M'_{mm}$	0	q'_m

Tabelle 2.2: Lineares Komplementaritätsproblem in Tableauform

Das Variablenpaar, daß vor dem letzten Pivotschritt Nichtbasisvariable war, wird im folgenden Schritt zur Basisvariablen. Die z -Komponente dieses Variablenpaars steht in Spalte k . Somit

wird die Spalte k zur Pivotspalte. Die neue Pivotzeile s wird nach der Zulässigkeitsregel

$$\frac{q_s}{M_{sk}} = \min_{M'_{jk} > 0} \left\{ \frac{q_j}{M_{jk}} \mid M_{jk} > 0 \right\} \quad (2.57)$$

bestimmt. Daraus folgt M_{sk} als Pivotelement und die mathematischen Umformungen können durch

$$q_s'' = -\frac{q_s'}{M'_{sk}}, \quad (2.58)$$

$$q_i'' = q_i - \frac{q_s'}{M'_{sk}} M'_{ik}, \quad i \neq s, \quad (2.59)$$

$$M''_{sj} = -\frac{M'_{sj}}{M'_{sk}}, \quad (2.60)$$

$$M''_{ij} = M_{ij} - \frac{M'_{sj}}{M'_{sk}} M'_{ik}, \quad i \neq s, \quad (2.61)$$

beschrieben werden.

In dieser Weise wird mit den Pivotschritten fortgefahren. Es wird immer die Pivotspalte zur Nichtbasisvariablen, deren Index mit der Variablen, die im letzten Pivotschritt Nichtbasisvariable wurde, übereinstimmt.

Das Verfahren endet mit einer von zwei Möglichkeiten:

1. Die zusätzliche Variable z_0 ist Nichtbasisvariable geworden. Damit ist die zugehörige Basislösung auch Lösung des LCP.
2. Falls die Zulässigkeitsregel nach Gleichung (2.61) nicht erfüllt werden kann, d.h. alle Zahlenwerte in der Pivotspalte sind negativ, kann keine Lösung des LCP gefunden werden.

2.4 Stochastische Grundlagen

Bei Anwendung des Kalman-Filters werden stochastische Signale betrachtet. Bei diesen Signalen handelt es sich um zeitabhängige Funktionen, deren Funktionswerte vom Zufall beeinflusst werden. In diesem Abschnitt wird auf die Begrifflichkeiten der Wahrscheinlichkeitsrechnung eingegangen, die zur Charakterisierung von zufälligen Signalen dienen. Eine ausführliche Erklärung und anschauliche Herleitung der Begriffe mit Beispielen ist in [7] zu finden.

Zunächst werden die fundamentalen Begriffe Elementarereignis und Zufallsexperiment anschaulich am Beispiel des Werfens eines Spielwürfels erklärt. Der Vorgang des Würfeln stellt das Zufallsexperiment dar, das Obenliegen einer Seite des Würfels, also das Ergebnis des Experimentes, wird als Elementarereignis bezeichnet. Die Menge, der sich paarweise gegenseitig ausschließenden Elementarereignisse, ist hier $\Omega = \{1, 2, 3, 4, 5, 6\}$. Die Wahrscheinlichkeit P des Eintretens eines zufälligen Ereignisses ist durch die Axiome von Kolmogorov definiert, die u.a. in [7] aufgeführt sind.

Zufallsvariable

Wenn Ω die Menge aller bei einem Zufallsexperiment möglichen Elementarereignisse beinhaltet, so kann jedes Ergebnis dieses Experimentes als Teilmenge von Ω dargestellt werden. Wird nun jedes Elementarereignis ω_i durch die Zuordnungsvorschrift ψ einem Vektor $\xi^{(i)}$ mit n reellen Zahlen zugeordnet, der eine Aussage über die Ausprägung des Ereignisses trifft, kann folgende mathematische Beschreibung getroffen werden

$$\xi^{(i)} = \psi(\omega_i) \quad , \omega_i \in \Omega. \quad (2.62)$$

Die Gleichung (2.62), die besagt, daß dem Elementarereignis ω_i die Zahlenwerte

$\xi^{(i)} = [\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(n)}]^T$ zugeordnet sind, kann durch die Gleichung

$$\mathbf{x} = \psi(\omega) \quad (2.63)$$

beschrieben werden. Tritt das Elementarereignis ω_i ein, nimmt die Variable \mathbf{x} den Wert $\xi^{(i)}$ an. Unter der Voraussetzung, daß zu unterschiedlichen ω_i auch verschiedene $\xi^{(i)}$ gehören, ist die Wahrscheinlichkeit für das Eintreten des Elementarereignisses ω_i gleich der Wahrscheinlichkeit, daß \mathbf{x} den Wert $\xi^{(i)}$ annimmt

$$P(\omega_i) \leq P(\mathbf{x} = \xi^{(i)}) \quad , \omega_i \in \Omega. \quad (2.64)$$

Der Vektor \mathbf{x} wird als vektorielle Zufallsvariable bezeichnet. Das „ \leq “-Zeichen in Gleichung (2.64) berücksichtigt die Tatsache, daß die Abbildung ψ nicht umkehrbar ist. Dies verdeutlicht, daß die Zufallsvariable \mathbf{x} für unterschiedliche Elementarereignisse ω_i die gleichen Zahlenwerte $\xi^{(i)}$ annehmen kann. Die umgekehrte Zuordnung ψ^{-1} ist nicht eindeutig. Jedem Elementarereignis ω_i sind also zwei Arten von Werten zugeordnet:

1. der Vektor $\xi^{(i)}$ der Zufallsvariable \mathbf{x} , der lediglich eine alternative Art der Bezeichnung von ω_i darstellt,
2. der Wert $P(\omega_i)$, der ein Maß für die Wahrscheinlichkeit des Ereignisses ist.

Es werden nun die charakteristischen Größen für stochastische Signale zunächst an Hand von skalaren Zufallsvariablen eingeführt. Im anschließenden Abschnitt werden die zuvor gefundenen Definitionen auf vektorielle Zufallsvariablen verallgemeinert.

2.4.1 Skalare Zufallsvariablen

Verteilungs- und Dichtefunktion: Da das zufällige Ereignis $(x \mid x < \xi)$ von ξ abhängig ist, ist auch die zugehörige Wahrscheinlichkeit $P(x < \xi)$ von ξ abhängig. Diese Abhängigkeit wird durch die Definitionsgleichung der Wahrscheinlichkeitsverteilungsfunktion $F(\xi)$ der skalaren Zufallsvariable x ausgedrückt

$$P(x < \xi) =: F(\xi). \quad (2.65)$$

Die Verteilungsfunktion $F(\xi)$ beschreibt die Wahrscheinlichkeit, daß die Zufallsvariable x einen Wert kleiner als ξ annimmt. Die Wahrscheinlichkeit, daß x innerhalb zweier Grenzen $\xi^{(a)}$ und $\xi^{(b)}$ liegt, kann ebenfalls mit Hilfe von $F(\xi)$ ausgedrückt werden

$$P(\xi^{(a)} \leq x < \xi^{(b)}) = F(\xi^{(b)}) - F(\xi^{(a)}). \quad (2.66)$$

Existiert die Ableitung $f(\xi) = \frac{\partial}{\partial \xi} F(\xi)$, so kann Gleichung (2.66) auch wie folgt ausgedrückt werden

$$\int_{\xi^{(a)}}^{\xi^{(b)}} f(\xi) d\xi = F(\xi^{(b)}) - F(\xi^{(a)}). \quad (2.67)$$

Es kann also neben der Verteilungsfunktion $F(\xi)$ auch mit deren Ableitung $f(\xi)$ die Wahrscheinlichkeit eines Ereignisses berechnet werden, das aus beliebigen zufälligen Elementarereignissen $\xi^{(a)} \leq x < \xi^{(b)}$ zusammengesetzt ist. $f(\xi)$ wird als die Dichtefunktion der eindimensionalen Zufallsvariable x bezeichnet.

Die Wahrscheinlichkeitsverteilung der Zufallsvariable x wird sowohl durch die Verteilungsfunktion $F(\xi)$ als auch durch die Dichtefunktion $f(\xi)$ beschrieben. Sie gibt darüber Aufschluß, welche Ereignisse häufig und welche weniger häufig eintreten.

Erwartungswert: Bei einem Zufallsexperiment seien m^* Elementarereignisse möglich, welche durch die Zuordnungsvorschrift ψ mit den Zahlenwerten ξ verknüpft sind. Es wird zugelassen, daß ψ nicht umkehrbar ist. Der Vektor ξ enthält deshalb m verschiedene Werte, wobei $m \leq m^*$ gilt. Ist die Zuordnung ψ umkehrbar, entspricht die Anzahl der Ereignisse der Zahlenwerte, d.h. es ist $m = m^*$. Diese den Elementarereignissen zugeordneten Zahlenwerte $\xi^{(i)}$ werden aufsteigend numeriert: $\xi = \{\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(m)}\}$. Wenn das Zufallsexperiment N -mal durchgeführt wird, wobei $N \gg m$ gilt, nimmt die Zufallsvariable x genau N_i -mal den Wert $\xi^{(i)}$ an. Das Ergebnis des Experiments ist

$$\begin{aligned} x = \xi^{(1)} & \quad \text{genau } N_1 \text{ mal,} \\ x = \xi^{(2)} & \quad \text{genau } N_2 \text{ mal,} \\ & \quad \vdots \\ x = \xi^{(m)} & \quad \text{genau } N_m \text{ mal.} \end{aligned}$$

Da die zufälligen Elementarereignisse $x = \xi^{(i)}$ sich paarweise gegenseitig ausschließen, also bei dem Eintreten des einen Ereignisses das andere nicht mehr geschehen kann, muß

$$\sum_{i=1}^m N_i = N \quad (2.68)$$

gelten. Für den arithmetischen Mittelwert aller $\xi^{(i)}$ ergibt sich

$$\bar{\xi} = \frac{1}{N} (\xi^{(1)} N_1 + \xi^{(2)} N_2 + \dots + \xi^{(m)} N_m) = \frac{1}{N} \sum_{i=1}^m \xi^{(i)} N_i = \sum_{i=1}^m \xi^{(i)} \frac{N_i}{N}. \quad (2.69)$$

Das Verhältnis $\frac{N_i}{N}$ beschreibt die relative Häufigkeit $h_i(N)$ des Auftretens von $\xi^{(i)}$. Für große Versuchszahlen stimmen $h_i(N)$ und die Wahrscheinlichkeit $P(x = \xi^{(i)})$ überein

$$h_i(N) \approx P(x = \xi^{(i)}) . \quad (2.70)$$

Gleichung (2.69) kann nun zu

$$\bar{\xi} = \sum_{i=1}^m \xi^{(i)} P(x = \xi^{(i)}) \quad (2.71)$$

formuliert werden. Nach Definition einer Wahrscheinlichkeitsfunktion $\tilde{f}(\xi)$ für eine diskrete Verteilung mit

$$\tilde{f}(\xi) := \begin{cases} P(x = \xi^{(i)}) & \text{für } \xi = \xi^{(i)} \quad i = 1, \dots, m \\ 0 & \text{sonst} \end{cases} \quad (2.72)$$

und Einsetzen dieses Ausdruckes in Gleichung (2.71), ergibt sich auf der rechten Seite die Definition des Erwartungswertes $E\{x\}$

$$E\{x\} := \sum_{i=1}^m \xi^{(i)} \tilde{f}(\xi^{(i)}) . \quad (2.73)$$

Wird an Stelle einer diskreten Wahrscheinlichkeitsverteilung eine stetige betrachtet, kann nach [7] mit der Dichtefunktion $f(\xi)$ die integrale Form des Erwartungswertes erhalten werden

$$E\{x\} = \int_{-\infty}^{+\infty} \xi f(\xi) d\xi . \quad (2.74)$$

Es sei besonders auf den Unterschied der beiden Werte $\bar{\xi}$ und $E\{x\}$ hingewiesen. Während der Mittelwert $\bar{\xi}$ vom jeweiligen Durchlauf des Zufallsexperimentes abhängig ist und dementsprechend variieren kann, ist der Erwartungswert eine charakteristische Größe für die Wahrscheinlichkeitsverteilung der Zufallsvariable x und somit ein fester Wert.

Varianz: Außer $\bar{\xi}$ kann noch eine andere Kennzahl aus den Versuchsergebnissen gebildet werden. Dies ist der quadratische Abstand der N Zahlenwerte der Zufallsvariablen x vom Mittelwert $\bar{\xi}$

$$\overline{(\xi - \bar{\xi})^2} = \frac{1}{N} \sum_{i=1}^m (\xi^{(i)} - \bar{\xi})^2 N_i = \sum_{i=1}^m (\xi^{(i)} - \bar{\xi})^2 \frac{N_i}{N} . \quad (2.75)$$

Analog zur Herleitung des Erwartungswertes $E\{x\}$ ergibt sich mit Gleichung (2.73)

$$\overline{(\xi - \bar{\xi})^2} = \sum_{i=1}^m (\xi^{(i)} - E\{x\})^2 \tilde{f}(\xi^{(i)}) . \quad (2.76)$$

Die rechte Seite der Gleichung (2.76) wird Varianz oder Streuung genannt und mit dem Symbol σ^2 bezeichnet. Die positive Quadratwurzel der Varianz heißt Standardabweichung.

Da die Varianz dem Mittelwert des Quadrates der Abweichung der Zufallsvariable x von ihrem Erwartungswert $E\{x\}$ entspricht, also der Größe $(x - E\{x\})^2$, kann sie auch als deren Erwartungswert geschrieben werden

$$\sigma^2 = E\left\{(x - E\{x\})^2\right\} = \sum_{i=1}^m (\xi^{(i)} - E\{x\})^2 f(\xi^{(i)}). \quad (2.77)$$

Ebenso wie der Erwartungswert kann auch die Varianz für eine stetige Wahrscheinlichkeitsfunktion dargestellt werden

$$\sigma^2 = E\left\{(x - E\{x\})^2\right\} = \int_{-\infty}^{+\infty} (\xi - E\{x\})^2 f(\xi) d\xi. \quad (2.78)$$

Nimmt x die Zahlenwerte ξ an, die weit um $E\{x\}$ streuen, ergibt sich ein großer Wert für die Varianz σ^2 . Liegen hingegen die Werte ξ dicht um den Erwartungswert verteilt, ist die Varianz klein. Die beiden Größen, Erwartungswert $E\{x\}$ und Varianz σ^2 , sind die Kennwerte der Wahrscheinlichkeitsverteilung. Als Beispiel sind für eine Gaußsche Normalverteilung der Erwartungswert $E\{x\}$ und die Standardabweichung σ in Abbildung 2.3 gezeigt.

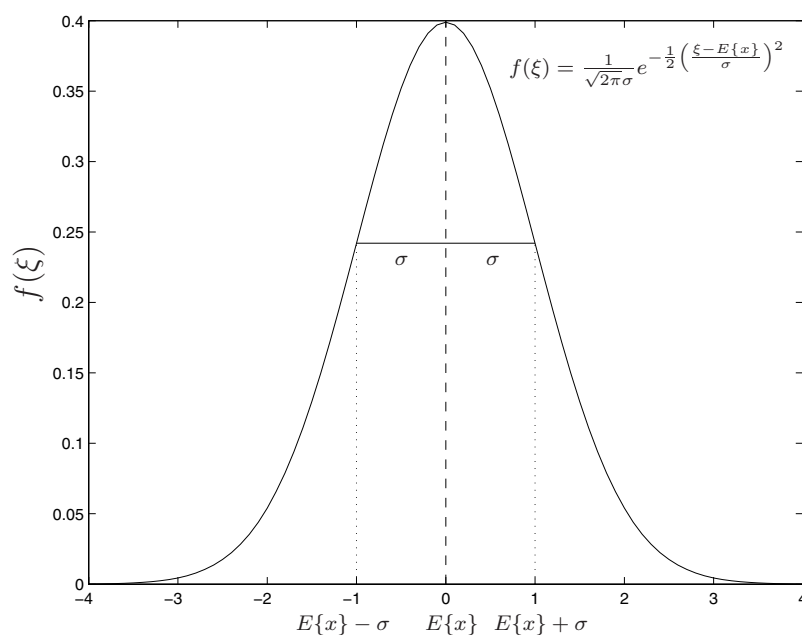


Abbildung 2.3: Dichtefunktion einer Gaußschen Normalverteilung

Momente: Sei $g(x)$ eine Funktion der Zufallsvariable x , dann kann ihr Erwartungswert $E\{g(x)\}$ mit Hilfe der Dichtefunktion $f(\xi)$ von x ausgedrückt werden

$$E\{g(x)\} = \int_{-\infty}^{+\infty} g(\xi) f(\xi) d\xi. \quad (2.79)$$

Die Erwartungswerte einer Funktion der Zufallsvariable x der Art $g(x) = x^k$ werden als Momente k -ter Ordnung bezeichnet

$$E\{x^k\} = \int_{-\infty}^{+\infty} \xi^k f(\xi) d\xi. \quad (2.80)$$

Allgemein wird das Zentralmoment k -ter Ordnung als der Erwartungswert einer Funktion $g(x) = (x - E\{x\})^k$ bezeichnet

$$E\{(x - E\{x\})^k\} = \int_{-\infty}^{+\infty} (\xi - E\{x\})^k f(\xi) d\xi. \quad (2.81)$$

Es ist ersichtlich, daß für skalare Zufallsvariablen das Moment 1. Ordnung dem Erwartungswert $E\{x\}$ der Zufallsvariable x entspricht und das Zentralmoment 2. Ordnung ihrer Varianz σ^2 .

2.4.2 Vektorielle Zufallsvariablen

Im vorherigen Abschnitt wurden nur eindimensionale Zufallsvariablen betrachtet, um die grundlegenden Begriffe möglichst einfach zu erläutern. Die meisten Zufallsprozesse besitzen in der Realität aber mehrere vom Zufall abhängige Größen, die zu einer vektoriellen Zufallsvariablen \mathbf{x} zusammengefaßt werden können. Die zuvor eingeführten Begriffe werden daher jetzt auf Vektoren erweitert. Der Übersicht halber werden zunächst nur zweidimensionale Zufallsvariablen $\mathbf{x} = [x_1, x_2]^T$ untersucht.

Verteilungs- und Dichtefunktion: Analog zu Gleichung (2.65) wird die Verteilungsfunktion einer zweidimensionalen Zufallsvariable $\mathbf{x} = [x_1, x_2]^T$ definiert

$$P((x_1 < \xi_1) \cap (x_2 < \xi_2)) =: F(\xi_1, \xi_2). \quad (2.82)$$

Ebenso kann die Dichtefunktion analog zu den Gleichungen (2.66) und (2.67)

$$P((\xi_1^{(a)} \leq x_1 < \xi_1^{(b)}) \cap (\xi_2^{(a)} \leq x_2 < \xi_2^{(b)})) = \int_{\xi_1^{(a)}}^{\xi_1^{(b)}} \int_{\xi_2^{(a)}}^{\xi_2^{(b)}} f(\xi_1, \xi_2) d\xi_1 d\xi_2 \quad (2.83)$$

geschrieben werden, falls die partielle Ableitung

$$f(\xi_1, \xi_2) = \frac{\partial^2}{\partial \xi_1 \partial \xi_2} F(\xi_1, \xi_2) \quad (2.84)$$

existiert.

Momente: Analog zum eindimensionalen Fall sind die Ausdrücke

$$E \{x_1^{k_1}, x_2^{k_2}\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \xi_1^{k_1} \xi_2^{k_2} f(\xi_1, \xi_2) d\xi_1 d\xi_2 \quad \text{und} \quad (2.85)$$

$$\begin{aligned} E \left\{ (x_1 - E\{x_1\})^{k_1}, (x_2 - E\{x_2\})^{k_2} \right\} \\ = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\xi_1 - E\{x_1\})^{k_1} (\xi_2 - E\{x_2\})^{k_2} f(\xi_1, \xi_2) d\xi_1 d\xi_2 \end{aligned} \quad (2.86)$$

als Moment bzw. Zentralmoment $(k_1 + k_2)$ -ter Ordnung definiert. Es existieren für zweidimensionale Zufallsvariablen zwei Momente 1. Ordnung. Eines für $k_1 = 1$ und $k_2 = 0$ sowie für $k_1 = 0$ und $k_2 = 1$. Diese beiden Momente sind die Erwartungswerte der skalaren Zufallsvariablen x_1 und x_2 , dementsprechend $E\{x_1\}$ und $E\{x_2\}$.

Für eine zweidimensionale Zufallsvariable sind drei Momente 2. Ordnung vorhanden. Gilt für das mittlere, also für jenes, das von x_1 als auch von x_2 abhängt, die Beziehung

$$E \{x_1, x_2\} = 0, \quad (2.87)$$

so werden x_1 und x_2 als orthogonal zu einander bezeichnet.

Wie bei den Momenten existieren ebenso drei Zentralmomente 2. Ordnung. Für den Fall, daß einer der beiden Exponenten k zu Null wird und damit der andere zu zwei, ergeben sich die Varianzen σ_1^2 und σ_2^2 . Das dritte Zentralmoment 2. Ordnung, bei dem k_1 und k_2 beide den Wert eins annehmen, wird als Kovarianz $cov(x_1, x_2)$ der Zufallsvariablen x_1 und x_2 bezeichnet

$$\begin{aligned} cov(x_1, x_2) &= E \left\{ (x_1 - E\{x_1\})(x_2 - E\{x_2\}) \right\} \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\xi_1 - E\{x_1\})(\xi_2 - E\{x_2\}) f(\xi_1, \xi_2) d\xi_1 d\xi_2 = cov(x_2, x_1). \end{aligned} \quad (2.88)$$

Genau genommen handelt es sich bei allen Zentralmomenten 2. Ordnung um Kovarianzen, da auch die Varianzen σ_1^2 und σ_2^2 Sonderfälle der Gleichung (2.88) sind:

$$cov(x_1, x_1) = E \left\{ (x_1 - E\{x_1\})^2 \right\} = \sigma_1^2 \quad (2.89)$$

$$cov(x_2, x_2) = E \left\{ (x_2 - E\{x_2\})^2 \right\} = \sigma_2^2 \quad (2.90)$$

Die Dichtefunktion $f(\xi_1, \xi_2)$ einer zweidimensionalen Zufallsvariable \mathbf{x} und damit ihre Wahrscheinlichkeitsverteilung sind nach [7] durch die zwei Momente 1. Ordnung sowie die drei Zentralmomente 2. Ordnung eindeutig festgelegt.

Zusätzlich wird noch der Korrelationskoeffizient ρ eingeführt, der die Stärke der statistischen Abhängigkeit der Zufallsvariablen x_1 und x_2 kennzeichnet. Er ist definiert durch

$$\rho = \frac{cov(x_1, x_2)}{\sigma_1 \sigma_2}. \quad (2.91)$$

Kann die Dichtefunktion $f(\xi_1, \xi_2)$ in ein Produkt aufgespalten werden, in dem die zwei Faktoren jeweils nur von ξ_1 bzw. ξ_2 abhängen, sind die zwei Zufallsvariablen x_1 und x_2 voneinander unabhängig. Dies bedeutet, daß die beiden Zufallsvariablen zwei stochastische Prozesse beschreiben, die sich gegenseitig nicht beeinflussen und getrennt voneinander untersucht werden können. Ist dies der Fall, so verschwindet ihre gemeinsame Kovarianz

$$\text{cov}(x_1, x_2) = E\left\{(x_1 - E\{x_1\})(x_2 - E\{x_2\})\right\} = 0. \quad (2.92)$$

Wie aus Gleichung (2.91) ersichtlich, wird der Korrelationskoeffizient $\rho = 0$. Somit sind x_1 und x_2 unkorreliert. Die Unkorreliertheit stellt eine schwächere Bedingung als die Unabhängigkeit dar. Erstere besagt, daß sich die statistischen Verteilungen der unkorrelierten Zufallsvariablen nicht gegenseitig beeinflussen. Über die Abhängigkeit der Zufallsvariablen an sich macht die Unkorreliertheit keine Aussage. Sind zwei Zufallsvariablen unabhängig voneinander, sind sie auch unkorreliert. Aus der Unkorreliertheit resultiert hingegen nicht ihre Unabhängigkeit.

Kovarianzmatrix: Der Erwartungswert einer von \mathbf{x} abhängigen $n \times m$ -Matrix $\mathbf{G}(x_1, x_2)$ ist definiert durch

$$\begin{aligned} E\left\{\mathbf{G}(x_1, x_2)\right\} &= E\left\{\begin{bmatrix} g_{11}(x_1, x_2) & \dots & g_{1m}(x_1, x_2) \\ \vdots & & \vdots \\ g_{n1}(x_1, x_2) & \dots & g_{nm}(x_1, x_2) \end{bmatrix}\right\} \\ &= \begin{bmatrix} E\{g_{11}(x_1, x_2)\} & \dots & E\{g_{1m}(x_1, x_2)\} \\ \vdots & & \vdots \\ E\{g_{n1}(x_1, x_2)\} & \dots & E\{g_{nm}(x_1, x_2)\} \end{bmatrix}. \end{aligned} \quad (2.93)$$

Mit der Gleichung (2.93) gelingt es alle Zentralmomente 2. Ordnung einer Zufallsvariable $\mathbf{x} = [x_1, x_2]^T$ in einem geschlossenen Ausdruck zusammenzufassen. Es wird dazu von dem dyadischen Produkt

$$\begin{aligned} (\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^T &= \begin{bmatrix} x_1 - E\{x_1\} \\ x_2 - E\{x_2\} \end{bmatrix} \begin{bmatrix} x_1 - E\{x_1\} & x_2 - E\{x_2\} \end{bmatrix} \\ &= \begin{bmatrix} (x_1 - E\{x_1\})^2 & (x_1 - E\{x_1\})(x_2 - E\{x_2\}) \\ (x_2 - E\{x_2\})(x_1 - E\{x_1\}) & (x_2 - E\{x_2\})^2 \end{bmatrix} \end{aligned} \quad (2.94)$$

der Erwartungswert gebildet

$$\begin{aligned} &E\left\{(\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^T\right\} \\ &= \begin{bmatrix} E\left\{(x_1 - E\{x_1\})^2\right\} & E\left\{(x_1 - E\{x_1\})(x_2 - E\{x_2\})\right\} \\ E\left\{(x_2 - E\{x_2\})(x_1 - E\{x_1\})\right\} & E\left\{(x_2 - E\{x_2\})^2\right\} \end{bmatrix}. \end{aligned} \quad (2.95)$$

Auf der Hauptdiagonalen dieser Matrix stehen die Varianzen der einzelnen Komponenten des Vektors \mathbf{x} . Die übrigen Matrixelemente sind die Kovarianzen $cov(x_1, x_2)$ und $cov(x_2, x_1)$, aus deren Identität die Symmetrie der Matrix resultiert. Aus den Gleichungen (2.89) und (2.90) wird deutlich, daß die Matrix ausschließlich Kovarianzen als Elemente enthält, so daß sie als Kovarianzmatrix $\mathbf{Cov}(\mathbf{x}, \mathbf{x})$ bezeichnet wird.

Bei Verallgemeinerung der zweidimensionalen Betrachtungen auf einen Vektor

$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ der Länge n , ergibt sich für die Kovarianzmatrix nach analoger Vorgehensweise der Ausdruck

$$\mathbf{Cov}(\mathbf{x}, \mathbf{x}) := E\left\{(\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^T\right\} = \begin{bmatrix} E\left\{(x_1 - E\{x_1\})^2\right\} & \dots & E\left\{(x_1 - E\{x_1\})(x_n - E\{x_n\})\right\} \\ \vdots & \ddots & \vdots \\ E\left\{(x_n - E\{x_n\})(x_1 - E\{x_1\})\right\} & \dots & E\left\{(x_n - E\{x_n\})^2\right\} \end{bmatrix}. \quad (2.96)$$

Falls der Erwartungswert $E\{\mathbf{x}\} = \mathbf{0}$ ist, vereinfacht sich die Kovarianzmatrix zu

$$\mathbf{Cov}(\mathbf{x}, \mathbf{x}) := E\{\mathbf{x}\mathbf{x}^T\} = \begin{bmatrix} E\{x_1^2\} & \dots & E\{x_1 x_n\} \\ \vdots & \ddots & \vdots \\ E\{x_n x_1\} & \dots & E\{x_n^2\} \end{bmatrix}. \quad (2.97)$$

Bei einer n -dimensionalen Zufallsvariable \mathbf{x} , deren Komponenten unabhängig und damit unkorreliert voneinander sind, verschwinden die Kovarianzen

$$cov(x_i, x_j) = 0 \quad \forall \quad i \neq j. \quad (2.98)$$

Damit wird $\mathbf{Cov}(\mathbf{x}, \mathbf{x})$ zu einer Diagonalmatrix. Sind bei zwei Zufallsvariablen \mathbf{x} und \mathbf{y} alle Komponenten von \mathbf{x} mit allen von \mathbf{y} unkorreliert, verschwindet die Kovarianzmatrix $\mathbf{Cov}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$.

3 Optimierungsmethoden

Zur Schadenserkennung werden Optimierungsstrategien aus unterschiedlichen Bereichen verwendet. Zum einen wird ein SQP-Verfahren getestet. Dieser Algorithmus repräsentiert die Verfahren aus dem Gebiet der Optimierung, die mittels des Gradienten der zu optimierenden Zielfunktion arbeiten. Das Kalman-Filter repräsentiert ein auf stochastischer Optimierung basierendes Verfahren. Es können innerhalb des Filters stochastische Werte des Optimierungsverlaufes berücksichtigt werden.

Aus dem Gebiet der evolutionären Strategien wird der genetische Algorithmus verwendet. Im wesentlichen ähnelt der Algorithmus der Evolutionsstrategie und des Simulated Annealing, weshalb auch nur ein Verfahren aus dieser Kategorie weiter untersucht wird.

Neuronale Netze werden als viertes Optimierungsverfahren zur Lösung des inversen Problems verwendet. Im Folgenden werden in zusammengefaßter Form die Grundlagen dieser Verfahren skizziert. Für eine ausführlichere Einführung in diese Themengebiete ist weitere Literatur angegeben.

3.1 SQP - Sequential Quadratic Programming

Die SQP-Verfahren sind zur Zeit die zuverlässigsten und effektivsten Verfahren der mathematischen Optimierung, die auf der Lösung der durch das Problem beschreibende Karush-Kuhn-Tucker Bedingungen basieren. SQP-Verfahren gelten als sehr nützlich für Optimierungsprobleme, für die keine Spezialverfahren entwickelt werden können. Im allgemeinen läßt sich die Optimierungsaufgabe mathematisch in der Form

$$\begin{aligned} \min f(\mathbf{x}) \\ g_j(\mathbf{x}) &= 0, \quad j = 1, \dots, m_e \\ g_j(\mathbf{x}) &\geq 0, \quad j = m_e + 1, \dots, m \\ x_l &\leq x \leq x_u \end{aligned} \tag{3.1}$$

beschreiben, in der $f(x)$ die zu minimierende Funktion unter den einzuhaltenden Nebenbedingungen $g_j(x)$ darstellt. Eine Lösung des im allgemeinen nichtlinearen Optimierungsproblems läßt sich durchführen, in dem die zugehörige Lagrange-Funktion betrachtet wird. Die Lagrange-

Funktion ist durch

$$L(\mathbf{x}, \mathbf{u}) = f(\mathbf{x}) - \sum_{j=1}^m u_j g_j(\mathbf{x}) \quad (3.2)$$

definiert, in der u_j die Lagrange-Multiplikatoren darstellen.

Das SQP-Verfahren verfolgt eine iterative Vorgehensweise zur Lösung des Optimierungsproblems (3.1). Demnach wird eine Näherung der Lösung nach der Iterationsvorschrift

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \mathbf{d} \quad (3.3)$$

bestimmt, wobei k die Anzahl der bisher erfolgten Iterationen angibt. In Gleichung (3.3) ist α eine positive Schrittweite und \mathbf{d} gibt die Richtung an, in der im folgenden Iterationsschritt $k+1$ der Wert für \mathbf{x}_k verändert werden soll. Innerhalb jedes Iterationsschrittes muß ein quadratisches Unterproblem der Form

$$\min f(\mathbf{x}_0) + \mathbf{g}^T(\mathbf{x} - \mathbf{x}_k) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_k)^T \mathbf{H}_k(\mathbf{x} - \mathbf{x}_k) \quad (3.4)$$

unter den Nebenbedingungen

$$g_j(\mathbf{x}_k) + \frac{\partial g_j(\mathbf{x}_k)}{\partial x_i} d_i = 0 \quad j = 1, \dots, m_e \quad (3.5)$$

$$g_j(\mathbf{x}_k) + \frac{\partial g_j(\mathbf{x}_k)}{\partial x_i} d_i \geq 0 \quad j = m_e + 1, \dots, m \quad (3.6)$$

gelöst werden, um die Suchrichtung \mathbf{d}_k zu erhalten. Dieses quadratische Unterproblem entsteht aus einer Näherung der Lagrange-Funktion (3.2) durch eine Taylorreihe im Punkt \mathbf{x}_k unter Vernachlässigung der Terme höherer Ordnung und der Linearisierung der Nebenbedingungen von (3.1). Dieses Optimierungsproblem wird gelöst, indem geeignete Löser quadratischer Optimierungsaufgaben eingesetzt werden. In (3.4) ist H_k die Hesse-Matrix oder eine der Hesse-Matrix angenäherte Matrix der Lagrange-Funktion (3.2). Desweiteren muß innerhalb jedes Iterationsschrittes k die Schrittweite α in Gleichung (3.3) neu bestimmt werden. Dies geschieht durch Minimierung der erweiterten Lagrange-Funktion

$$L(\mathbf{x}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) - \sum_{i=m_e+1}^m u_i (g_i(\mathbf{x} - s_i) + \frac{1}{2} \sum_{i=m_e+1}^m \rho_i (g_i(\mathbf{x}) - s_i)^2) \quad (3.7)$$

als Meritfunktion. Nach Abschluß dieser Operation wird \mathbf{x}_{k+1} in Gleichung (3.3) bestimmt. Zur Berechnung der Näherungslösung für die Hesse-Matrix der Lagrange-Funktion wird das Broyden-Fletcher-Goldfarb-Shanno (BFGS) Verfahren mit

$$\mathbf{H}_{k+1} = \mathbf{H}_k - \frac{1}{\Delta \mathbf{x}^T \mathbf{H}_k \Delta \mathbf{x}} \mathbf{H}_k \Delta \mathbf{x} \Delta \mathbf{x}^T \mathbf{H}_k + \frac{1}{\mathbf{y}^T \Delta \mathbf{x}} \mathbf{y} \mathbf{y}^T \quad (3.8)$$

benutzt, in der $\Delta \mathbf{x} = \mathbf{x}_{k+1} - \mathbf{x}_k$ und \mathbf{y} die Änderung des Gradienten der zu minimierenden Funktion f mit

$$y_{ik} = \frac{\partial f(\mathbf{x}_{k+1})}{\partial x_i} - \frac{\partial f(\mathbf{x}_k)}{\partial x_i} \quad i = 1, \dots, m \quad (3.9)$$

angibt.

Für die numerischen Untersuchungen wird das SQP-Verfahren aus der NAG-Library angewendet. SQP-Methoden sind ausführlich in [15] und [39] erklärt.

3.2 Kalman-Filter

3.2.1 Stochastischer Signalprozeß

In Kapitel 2.4 wurde auf grundlegende Begriffe aus der Wahrscheinlichkeitsrechnung eingegangen, die für den Umgang mit stochastischen Signalprozeßen notwendig sind. Es wurden dabei allerdings keine Angaben darüber gemacht, was stochastische Signalprozeße charakterisiert und wo sie Anwendung finden. Dies soll in diesem Abschnitt geschehen.

Ein zu regelnder Prozeß beinhaltet in der Realität meistens regellose, also stochastische Störsignale, über deren Verlauf in der Zeit keine deterministische Beschreibung berechnet werden kann. Um den Aufwand für den Entwurf einer Regelung gering zu halten, können deterministische Störsignale verwendet werden, also Signale, die im voraus exakt bekannt sind und analytisch beschrieben werden können. Diese deterministischen Störsignale sind dann nur „Ersatzsignale“ der wirklich vorkommenden. Die deterministische Prozeßregelung ist natürlich nur optimal bezüglich der Ersatzsignale.

Werden höhere Anforderungen an die Qualität einer Regelung gestellt, ist es notwendig neben dem dynamischen Verhalten des Prozesses auch die Einwirkungen der Störgrößen der Wirklichkeit besser anzupassen. Dies kann durch die Theorie der stochastischen Störsignale geschehen.

Grundbegriffe

Ein dynamischer Prozeß, dessen Parameter stochastische Größen sind, wird als stochastischer Signalprozeß bezeichnet. Betrachtet man nach [22] einen zeitdiskreten, stochastischen, skalaren Signalprozeß $x(t_k)$, bei dem t_k die diskreten Zeitpunkte darstellt, werden seine statistischen Eigenschaften durch die Wahrscheinlichkeitsverteilungsfunktion bzw. die Dichtefunktion seiner Signale beschrieben. Sind die Verteilungsdichten Funktionen der Zeit, wird das Signal nach [22] als nichtstationär bezeichnet. Wenn die Dichtefunktionen aber unabhängig von einer Zeitverschiebung sind, ist das stochastische Signal stationär im strengen Sinn.

Der Erwartungswert $E\{x(t_k)\}$ zu einem festen Zeitpunkt t_k wird aus der Mittelung aller Funktionswerte $\xi^{(1)}(t), \xi^{(2)}(t), \dots, \xi^{(m)}(t)$ zum gleichen Zeitpunkt t_k gewonnen. Diese Mittelung wird Ensemblemittelung genannt. Die Mittelung eines einzigen Funktionswertes $\xi^{(i)}$ über die gesamte Zeit ist davon streng zu unterscheiden.

Kann bei einem stationären Signal die Ensemblemittelung durch die Zeitmittelung ersetzt werden, wird das Signal ergodisch genannt. Die Signale können dann durch ihren Erwartungswert $E\{x(t_k)\}$ und die Autokorrelationsfunktion $\Phi_{xx}(\tau)$ beschrieben werden. Die Autokorrelations-

funktion

$$\Phi_{xx}(\tau) = E\{x(t_k) x(t_k + \tau)\} \quad (3.10)$$

ist ein Maß für die inneren Zusammenhänge eines regellosen Signals. Mit $\Phi_{xx}(\tau)$ wird eine Aussage über das wahrscheinliche Verhalten des Signals $x(t)$ in seinem zeitlichen Verlauf gemacht. Die gleiche Aussagekraft wie die Autokorrelation hat die Autokovarianzfunktion $\text{cov}(x, \tau)$.

$$\text{cov}(x, \tau) = E\{(x(t_k) - \bar{x})(x(t_k + \tau) - \bar{x})\} = \Phi_{xx}(\tau) - \bar{x}^2. \quad (3.11)$$

Bei einer Gaußschen Wahrscheinlichkeitsverteilung wird das Signal durch seinen Erwartungswert $E\{x\}$ und seine Autokovarianz $\text{cov}(x, \tau)$ vollständig beschrieben. In Gleichung 3.11 stellt \bar{x} den Mittelwert der Folge $x(t_k)$ dar.

Ein stochastisches Signal wird nach [7] stationär im weiten Sinn genannt, wenn sein Erwartungswert $E\{x\}$ zeitunabhängig ist, und seine Autokorrelationsfunktion $\Phi_{xx}(\tau)$ nur von der Zeitverschiebung τ abhängt.

Für zwei verschiedene stationäre, stochastische Signale ist ihre gegenseitige Abhängigkeit durch die Kreuzkorrelationsfunktion bzw. die Kreuzkovarianzfunktion beschrieben

$$\Phi_{xy}(\tau) = E\{x(t_k)y(t_k + \tau)\}, \quad (3.12)$$

$$\begin{aligned} \text{cov}(x, y, \tau) &= E\{(x(t_k) - E\{x\})(y(t_k + \tau) - E\{y\})\} \\ &= \Phi_{xy}(\tau) - E\{x\}E\{y\}. \end{aligned} \quad (3.13)$$

Wird die Kreuzkovarianzfunktion zweier Signale zu Null,

$$\text{cov}(x, y, \tau) = 0, \quad (3.14)$$

werden x und y unkorreliert genannt. Ihre Wahrscheinlichkeitsverteilungen beeinflussen sich also gegenseitig nicht. Sie sind orthogonal, falls zusätzlich noch

$$E\{x\}E\{y\} = 0 \Rightarrow \Phi_{xy} = 0 \quad (3.15)$$

erfüllt ist.

Die ungewollten Anteile eines Signals werden als Rauschen bezeichnet. So genanntes weißes Rauschen ist dadurch gekennzeichnet, daß der momentane Signalwert unabhängig von allen vergangenen Werten ist, d.h. weißes Rauschen besitzt keine inneren Zusammenhänge. Es ist durch seinen Erwartungswert und seine Kovarianz

$$\text{cov}(x, \tau) = \begin{cases} \sigma_x & \tau = 0 \\ 0 & \tau \neq 0 \end{cases} \quad (3.16)$$

bei Gaußscher Normalverteilung vollständig beschrieben.

Die bisherigen Betrachtungen lassen sich von skalaren auf vektorielle Signale erweitern. Der innere Zusammenhang eines vektoriellen Signals

$$\mathbf{x}(t_k) = [x_1(t_k) x_2(t_k) \dots x_n(t_k)]^T \quad (3.17)$$

wird durch die Kovarianzmatrix zum Ausdruck gebracht. Auf ihrer Hauptdiagonalen stehen die Autokovarianzen, alle anderen Elemente sind die jeweiligen Kreuzkovarianzen zweier Komponenten des Signalvektors. Für ein n -dimensionales Signal ergibt sich die folgende Kovarianzmatrix

$$\mathbf{Cov}(\mathbf{x}(t_k), \tau) = \begin{bmatrix} \Phi_{x_1 x_1}(\tau) - E\{x_1\}^2 & \dots & \Phi_{x_1 x_n}(\tau) - E\{x_1\}E\{x_n\} \\ \vdots & \ddots & \vdots \\ \Phi_{x_n x_1}(\tau) - E\{x_n\}E\{x_1\} & \dots & \Phi_{x_n x_n}(\tau) - E\{x_n\}^2 \end{bmatrix}. \quad (3.18)$$

Für $\tau = 0$ ergibt sich die aus Gleichung (2.96) bekannte symmetrische Kovarianzmatrix.

3.2.2 Lineare Kalman-Filterung in diskreter Zeit

Das Kalman-Filter dient zur Schätzung von Zuständen eines Systems, die nicht unmittelbar zugänglich sind. Die Schätzung wird rekursiv durchgeführt, d.h. die aktuellen Schätzwerte werden durch eine Verbesserung der unmittelbar vorangegangenen Werte erzeugt. Die zu schätzenden Zustandsgrößen werden durch ein dynamisches, zeitdiskretes, lineares, stochastisches Modell beschrieben. In diesem Abschnitt wird das lineare Kalman-Filter hergeleitet, wie es z.B. in [6], [8], [10], [16] oder [51] zu finden ist.

Zunächst wird eine abkürzende Schreibweise eingeführt, in der die Abhängigkeit einer Größe vom Zeitpunkt t_k durch den tiefgestellten Index $_k$ gekennzeichnet ist. Es folgt die Auflistung der Symbole und Benennungen der hier verwendeten Größen zum Zeitpunkt t_k :

\mathbf{x}_k	: Vektor der Zustandsgrößen	$[n \times 1]$
$\hat{\mathbf{x}}_k$: Vektor der Schätzwerte der Zustandsgrößen	$[n \times 1]$
Φ_k	: Transitions- oder Systemmatrix des diskreten Prozesses	$[n \times n]$
\mathbf{z}_k	: Vektor der Meßwerte	$[m \times 1]$
\mathbf{H}_k	: Meßmatrix	$[m \times n]$
\mathbf{w}_k	: Vektor der Störgrößen	$[n \times 1]$
\mathbf{Q}_k	: Kovarianzmatrix der Störgrößen $\mathbf{w}(k)$	$[n \times n]$
\mathbf{v}_k	: Vektor der Meßfehler	$[m \times 1]$
\mathbf{R}_k	: Kovarianzmatrix der Meßfehler $\mathbf{v}(k)$	$[m \times m]$
\mathbf{e}_k	: Vektor der Schätzfehler	$[n \times 1]$
\mathbf{P}_k	: Kovarianzmatrix der Schätzfehler \mathbf{e}_k	$[n \times n]$
\mathbf{K}_k	: Kalman-Verstärkungsmatrix	$[n \times m]$

Der im Folgenden verwendete hochgestellte Index $-$ kennzeichnet die aus dem vorangegangenen Zeitschritt t_{k-1} extrapolierten Größen, die noch nicht mit Hilfe der im Zeitschritt t_k beobachteten Größen verbessert wurden. Hinter den jeweiligen Bezeichnungen der Matrizen bzw. Vektoren ist deren Dimension angegeben, wobei n die Anzahl der Zustandsgrößen und m die der Meßwerte symbolisiert.

Allgemeines Kalman-Filter

Ein stochastischer Signalprozeß kann durch die Systemgleichung

$$\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \mathbf{w}_k \quad (3.19)$$

dargestellt werden. Die Transitionsmatrix Φ_k beschreibt den Übergang des Zustandsvektors \mathbf{x}_k vom Zeitpunkt t_k nach t_{k+1} . Die Beobachtung dieses Prozesses an diskreten Zeitpunkten kann durch die lineare Beziehung

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (3.20)$$

ausgedrückt werden.

Die Störgrößen \mathbf{w}_k und die Meßfehler \mathbf{v}_k werden als weißes Rauschen und gegenseitig voneinander unabhängig angenommen. Ihre Kovarianzmatrizen seien gegeben mit

$$E \{ \mathbf{w}_k \mathbf{w}_i^T \} = \begin{cases} \mathbf{Q}_k, & i = k \\ 0, & i \neq k \end{cases}, \quad (3.21)$$

$$E \{ \mathbf{v}_k \mathbf{v}_i^T \} = \begin{cases} \mathbf{R}_k, & i = k \\ 0, & i \neq k \end{cases}, \quad (3.22)$$

$$E \{ \mathbf{w}_k \mathbf{v}_i^T \} = 0, \quad \forall k, i. \quad (3.23)$$

Während eines Schrittes der rekursiven Schätzung wird zunächst der vorangegangene, geschätzte Zustandsvektor $\hat{\mathbf{x}}_{k-1}$ mit der Systemgleichung (3.19) in die aktuelle Zeit überführt. Durch $\hat{\mathbf{x}}_k^-$ sei diese a priori Schätzung symbolisiert, also jene, die noch nicht durch die neuen Informationen der Meßwerte aktualisiert wurde. Der a priori Schätzfehler berechnet sich zu

$$\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^-. \quad (3.24)$$

Bei Annahme, daß der Schätzfehler \mathbf{e}_k^- den Erwartungswert $E \{ \mathbf{e}_k^- \} = \mathbf{0}$ hat, berechnet sich die Fehlerkovarianzmatrix nach Gleichung (2.97) zu

$$\mathbf{P}_k^- = E \{ \mathbf{e}_k^- \mathbf{e}_k^{-T} \} = E \{ (\mathbf{x}_k - \hat{\mathbf{x}}_k^-)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)^T \}. \quad (3.25)$$

Die Messungen \mathbf{z}_k sollen dazu benutzt werden, den a priori Schätzwert $\hat{\mathbf{x}}_k^-$ zu verbessern. Dazu wird die Methode der linearen Wichtung des verrauschten Meßfehlers $(\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-)$ und der a priori Schätzung $\hat{\mathbf{x}}_k^-$ gewählt.

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-). \quad (3.26)$$

Der durch die Meßwerte aktualisierte a posteriori Schätzwert wird mit $\hat{\mathbf{x}}_k$ bezeichnet. Der zugehörige Schätzfehler ist analog zu Gleichung (3.24)

$$\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k. \quad (3.27)$$

Das Modell des beobachteten Systems und der Schätzung durch das Kalman-Filter sind in Abbildung 3.1 dargestellt.

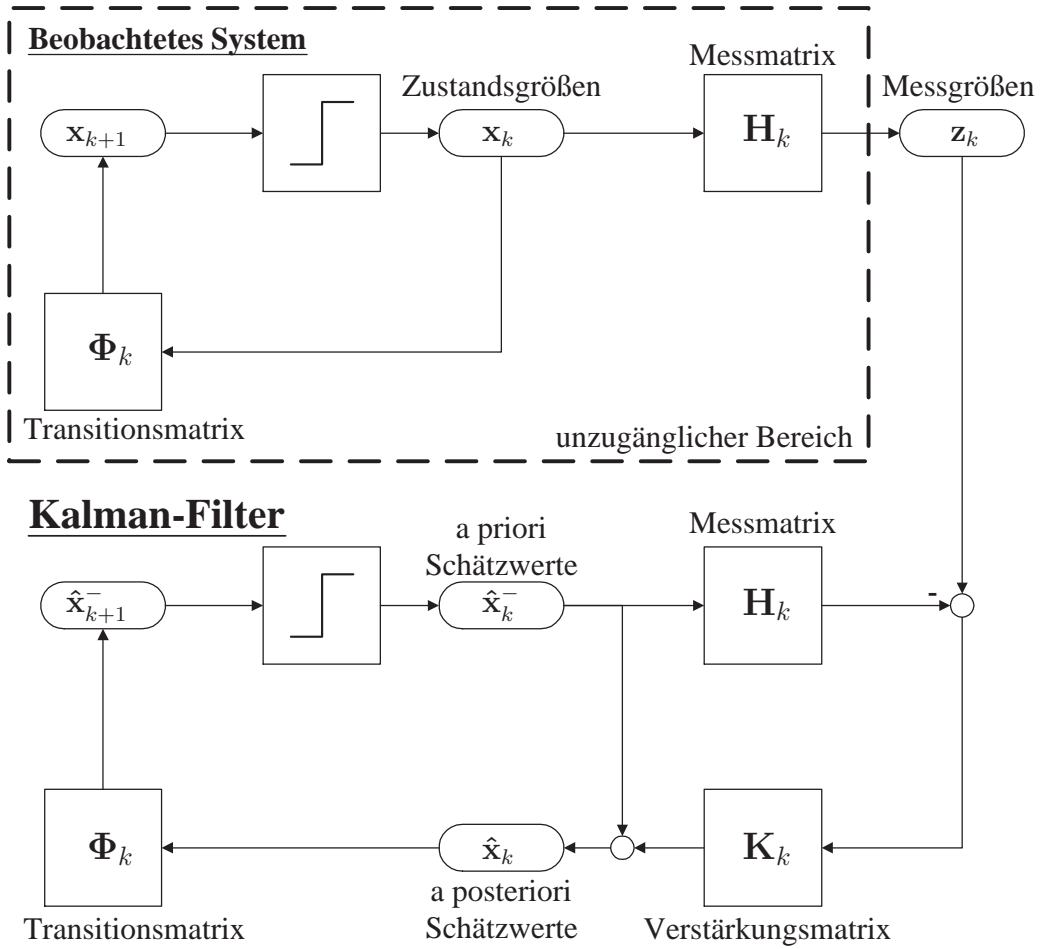


Abbildung 3.1: Modell des beobachteten Systems und der Schätzung durch das Kalman-Filter

\mathbf{K}_k ist der noch unbekannte Wichtungsfaktor, der so zu bestimmen ist, daß sich eine optimale Schätzung ergibt. Hierfür wird zunächst die a posteriori Fehlerkovarianzmatrix ermittelt

$$\mathbf{P}_k = E \{ \mathbf{e}_k \mathbf{e}_k^T \} = E \{ (\mathbf{x}_k - \hat{\mathbf{x}}_k) (\mathbf{x}_k - \hat{\mathbf{x}}_k)^T \}. \quad (3.28)$$

Einsetzen der Gleichungen (3.26) und (3.20) in die Gleichung (3.28) ergibt folgenden Zusammenhang

$$\mathbf{P}_k = E \left\{ \left(\mathbf{x}_k - \hat{\mathbf{x}}_k^- - \mathbf{K}_k (\mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \right) \left(\mathbf{x}_k - \hat{\mathbf{x}}_k^- - \mathbf{K}_k (\mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \right)^T \right\}. \quad (3.29)$$

Da der a priori Schätzfehler \mathbf{e}_k^- und der Meßfehler \mathbf{v}_k unkorreliert sind, ist es möglich ihren gemeinsamen Erwartungswert als Produkt der einzelnen Zufallsgrößen zu schreiben. Es wird

dadurch mit Gleichung (3.24) folgende Umformung möglich:

$$\begin{aligned} \mathbf{P}_k &= E \left\{ \begin{aligned} &(\mathbf{x}_k - \hat{\mathbf{x}}_k^- - \mathbf{K}_k \mathbf{v}_k - \mathbf{K}_k \mathbf{H}_k (\mathbf{x}_k - \hat{\mathbf{x}}_k^-)) \\ &(\mathbf{x}_k - \hat{\mathbf{x}}_k^- - \mathbf{K}_k \mathbf{v}_k - \mathbf{K}_k \mathbf{H}_k (\mathbf{x}_k - \hat{\mathbf{x}}_k^-))^T \end{aligned} \right\} \\ &= E \left\{ \begin{aligned} &((\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)) ((\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-))^T \\ &- \mathbf{K}_k E \{ \mathbf{v}_k \mathbf{v}_k^T \} \mathbf{K}_k^T. \end{aligned} \right\} \end{aligned} \quad (3.30)$$

Somit ergibt sich für \mathbf{P}_k folgende Beziehung

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) E \{ \mathbf{e}_k^- \mathbf{e}_k^{-T} \} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T - \mathbf{K}_k E \{ \mathbf{v}_k \mathbf{v}_k^T \} \mathbf{K}_k^T. \quad (3.31)$$

Unter der Berücksichtigung der Gleichungen (3.22) und (3.25) ergibt sich daraus

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^- (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T. \quad (3.32)$$

Eine optimale Schätzung kann erhalten werden, in dem ein \mathbf{K}_k bestimmt wird, das die individuellen Terme auf der Hauptdiagonalen der Fehlerkovarianzmatrix \mathbf{P}_k minimal werden läßt. Diese Terme repräsentieren die Fehlervarianzen $\sigma_{e_{ki}}^2 = \text{cov}(e_{ki} e_{ki})$ des zu schätzenden Zustandsvektors \mathbf{x}_k .

Nach Ausmultiplizieren und Umformen der Gleichung (3.32) entsteht die Gleichung

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_k^- - \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T. \quad (3.33)$$

Diese Gleichung ist quadratisch bezüglich \mathbf{K}_k . Jetzt soll die Summe der Fehlerquadrate der Zustandsgrößen minimiert werden, die der Spur der Matrix \mathbf{P}_k entspricht. Wird Gleichung 3.33 differenziert unter Berücksichtigung der Tatsache, daß die Spur von $\mathbf{K}_k \mathbf{H}_k \mathbf{P}_k^-$ der von $\mathbf{P}_k^- \mathbf{H}_k^T \mathbf{K}_k^T$ entspricht, entsteht der Ausdruck

$$\frac{d(\text{tr } \mathbf{P}_k)}{d\mathbf{K}_k} = -2 (\mathbf{H}_k \mathbf{P}_k^-)^T + 2\mathbf{K}_k (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k) \stackrel{!}{=} \mathbf{0}. \quad (3.34)$$

Hieraus ergibt sich der optimale Wichtungsfaktor, der die Streuung der Fehlerquadrate des Schätzfehlers minimiert

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1}. \quad (3.35)$$

Die entstandene Matrix \mathbf{K}_k wird Kalman-Verstärkungsmatrix genannt.

Um die Fehlerkovarianzmatrix \mathbf{P}_k zu erhalten, die der optimalen Schätzung zugeordnet ist, wird die Gleichung (3.35) in (3.33) eingesetzt

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \mathbf{H}_k \mathbf{P}_k^- \quad (3.36)$$

$$\Leftrightarrow \mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T \quad (3.37)$$

$$\Leftrightarrow \mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^-. \quad (3.38)$$

Es stehen jetzt vier Gleichungen zur Berechnung von \mathbf{P}_k zur Verfügung, wobei die Gleichungen (3.36), (3.37) und (3.38) im Unterschied zur Gleichung (3.32) nur dann gültig sind, falls \mathbf{K}_k die

optimale Verstärkungsmatrix darstellt. Da das Kalman-Filter in der Anwendung eine numerische Prozedur ist, ist die Wahl der günstigsten Darstellungsform von der konkret zu lösenden Aufgabenstellung abhängig.

Als letzte Rechnung des rekursiven Iterationsschrittes erfolgt nun noch analog zu Gleichung (3.19) die Extrapolation des Schätzwertes $\hat{\mathbf{x}}_k$ aus dem Zeitschritt t_k nach t_{k+1} durch die Transitionsmatrix Φ_k . Die Störgröße \mathbf{w}_k stellt ein weißes Rauschen dar, d.h. sie ist mit den vorherigen Störgrößen unkorreliert und ihr Erwartungswert ist dementsprechend $E\{\mathbf{w}_k\} = \mathbf{0}$. Der Vektor \mathbf{w}_k kann daher für die Transition ignoriert werden, da er stochastisch gesehen verschwindet

$$\hat{\mathbf{x}}_{k+1}^- = \Phi_k \hat{\mathbf{x}}_k. \quad (3.39)$$

Für die Bestimmung der mit der extrapolierten geschätzten Zustandsgröße $\hat{\mathbf{x}}_{k+1}^-$ assoziierten Fehlerkovarianzmatrix \mathbf{P}_{k+1}^- zum Zeitpunkt t_{k+1} wird zunächst der Fehler \mathbf{e}_{k+1}^- durch die Gleichungen (3.19), (3.27) und (3.39) ausgedrückt

$$\begin{aligned} \mathbf{e}_{k+1}^- &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}^- \\ &= (\Phi_k \mathbf{x}_k + \mathbf{w}_k) - \Phi_k \hat{\mathbf{x}}_k \\ &= \Phi_k \mathbf{e}_k + \mathbf{w}_k. \end{aligned} \quad (3.40)$$

Da \mathbf{w}_k das Prozeßrauschen zur Bestimmung der Zustandsgröße \mathbf{x}_{k+1} symbolisiert, also erst jenes für den folgenden Zeitpunkt t_{k+1} , sind \mathbf{e}_k und \mathbf{w}_k unkorreliert. Daher ergibt sich mit Gleichung (3.40) für die Matrix \mathbf{P}_{k+1}^-

$$\begin{aligned} \mathbf{P}_{k+1}^- &= E\left\{\mathbf{e}_{k+1}^- \mathbf{e}_{k+1}^{-T}\right\} = E\left\{(\Phi_k \mathbf{e}_k + \mathbf{w}_k)(\Phi_k \mathbf{e}_k + \mathbf{w}_k)^T\right\} \\ &= \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k. \end{aligned} \quad (3.41)$$

3.2.3 Regulärer Kalman-Filteralgorithmus

Für den zuvor hergeleiteten rekursiven Algorithmus, der als reguläre Form der Kalman-Filterung bezeichnet wird, seien nochmals die einzelnen Schritte und charakteristischen Gleichungen in der Reihenfolge ihrer Berechnung angeführt.

0.) Voraussetzungen:

- a) Φ_k und \mathbf{H}_k müssen exakt bekannt sein.
- b) Die Anfangswerte $\hat{\mathbf{x}}_0^-$ und \mathbf{P}_0^- müssen bekannt sein oder geeignet angenommen werden.

1.) Kalman-Verstärkungsmatrix zur Korrektur der Vorhersage $\hat{\mathbf{x}}_k^-$ durch neue Messungen \mathbf{z}_k

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (3.42)$$

- 2.) Korrektur der Vorhersage durch gewichtete Mittelwertbildung aus $\hat{\mathbf{x}}_k^-$ und \mathbf{z}_k

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \quad (3.43)$$

- 3.) Kovarianzmatrix (a posteriori) des Fehlers \mathbf{e}_k des korrigierten Schätzwertes $\hat{\mathbf{x}}_k$ berechnen

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^- \quad (3.44)$$

- 4.) Extrapolation des Schätzwertes in die Zeit t_{k+1}

$$\hat{\mathbf{x}}_{k+1}^- = \Phi_k \hat{\mathbf{x}}_k \quad (3.45)$$

- 5.) Kovarianzmatrix (a priori) des Fehlers \mathbf{e}_k^- des vorhergesagten Schätzwertes $\hat{\mathbf{x}}_{k+1}^-$ berechnen

$$\mathbf{P}_{k+1}^- = \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k \quad (3.46)$$

- 6.) Erhöhe k um Eins und gehe zu 1.)

Die Schleife des Filteralgorithmus ist in Abbildung 3.2 dargestellt.

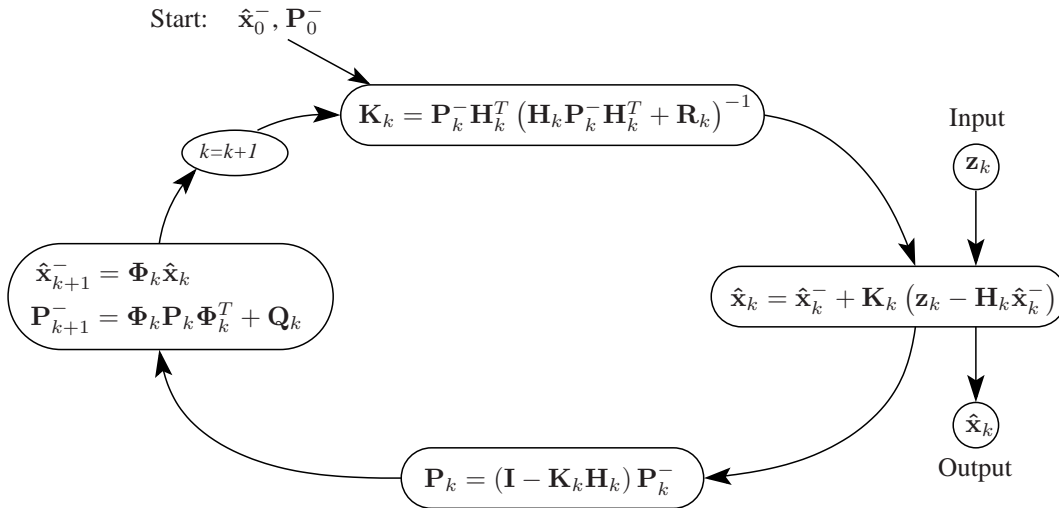


Abbildung 3.2: Regulärer Kalman-Filteralgorithmus

Alternative Form der Kalman-Filterung

Die Gleichungen des regulären Kalman-Filters können algebraisch manipuliert werden, um eine für bestimmte Anwendungsfälle numerisch günstigere Form zu erhalten. Zunächst werden hier diese Manipulationen nach [8] vorgeführt, danach wird auf die Arten der Problemstellungen eingegangen, die sich durch diese alternative Form des Kalman-Filters besser lösen lassen. Ausgegangen wird von den Gleichungen (3.42) bis (3.46) des regulären Kalman-Filters. Das Einsetzen von Gleichung (3.42) in (3.44) führt zu dem schon oben in Gleichung (3.36) erwähnten Ausdruck für die Fehlerkovarianzmatrix

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \mathbf{H}_k \mathbf{P}_k^- \quad (3.47)$$

Existieren die Inversen von \mathbf{P}_k , \mathbf{P}_k^- und \mathbf{R}_k , kann \mathbf{P}_k^{-1} , wie in [8] gezeigt, folgendermaßen ausgedrückt werden

$$\mathbf{P}_k^{-1} = (\mathbf{P}_k^-)^{-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k. \quad (3.48)$$

Es kann ebenfalls für die Kalman-Verstärkungsmatrix \mathbf{K}_k eine alternative Ausdrucksform gefunden werden. Dazu wird zunächst Gleichung (3.42) mit den Einheitsmatrizen $\mathbf{P}_k \mathbf{P}_k^{-1}$ und $\mathbf{R}_k^{-1} \mathbf{R}_k$ multipliziert.

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_k \mathbf{P}_k^{-1} \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{R}_k (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \\ &= \mathbf{P}_k \mathbf{P}_k^{-1} \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} + \mathbf{I})^{-1} \end{aligned} \quad (3.49)$$

Einsetzen der Gleichung (3.48) in den Ausdruck (3.49) liefert

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_k \left((\mathbf{P}_k^-)^{-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \right) \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} + \mathbf{I})^{-1} \\ &= \mathbf{P}_k (\mathbf{I} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{P}_k^-) \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} + \mathbf{I})^{-1} \\ &= \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{I} + \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1}) (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{R}_k^{-1} + \mathbf{I})^{-1} \\ &= \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1}. \end{aligned} \quad (3.50)$$

Der alternative Kalman-Filteralgorithmus ergibt sich durch das Substituieren der Gleichungen (3.42) und (3.44) für die Bestimmung der Kalman-Verstärkungsmatrix und der aktualisierten Fehlerkovarianzmatrix mit den Gleichungen (3.50) und (3.48).

Es ist hierbei zu bemerken, daß die aktualisierte Fehlerkovarianzmatrix \mathbf{P}_k unabhängig von der Kalman-Verstärkungsmatrix \mathbf{K}_k bestimmt werden kann. Andererseits muß für die Berechnung von \mathbf{K}_k die Kenntnis von \mathbf{P}_k vorausgesetzt werden. Hieraus ergibt sich eine Änderung der Reihenfolge der Gleichungen des Filteralgorithmus in alternativer Form gegenüber der regulären.

Alternativer Kalman-Filteralgorithmus

Die fünf Gleichungen der Alternativen Kalman-Filter Sequenz sind nachfolgend in ihrer Reihenfolge aufgeführt:

- 1.) $\mathbf{P}_k^{-1} = (\mathbf{P}_k^-)^{-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k$
- 2.) $\mathbf{K}_k = \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1}$
- 3.) $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-)$
- 4.) $\hat{\mathbf{x}}_{k+1}^- = \Phi_k \hat{\mathbf{x}}_k$
- 5.) $\mathbf{P}_{k+1}^- = \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k$

In Abbildung 3.3 ist die Schleife des alternativen Kalman-Filteralgorithmus dargestellt.

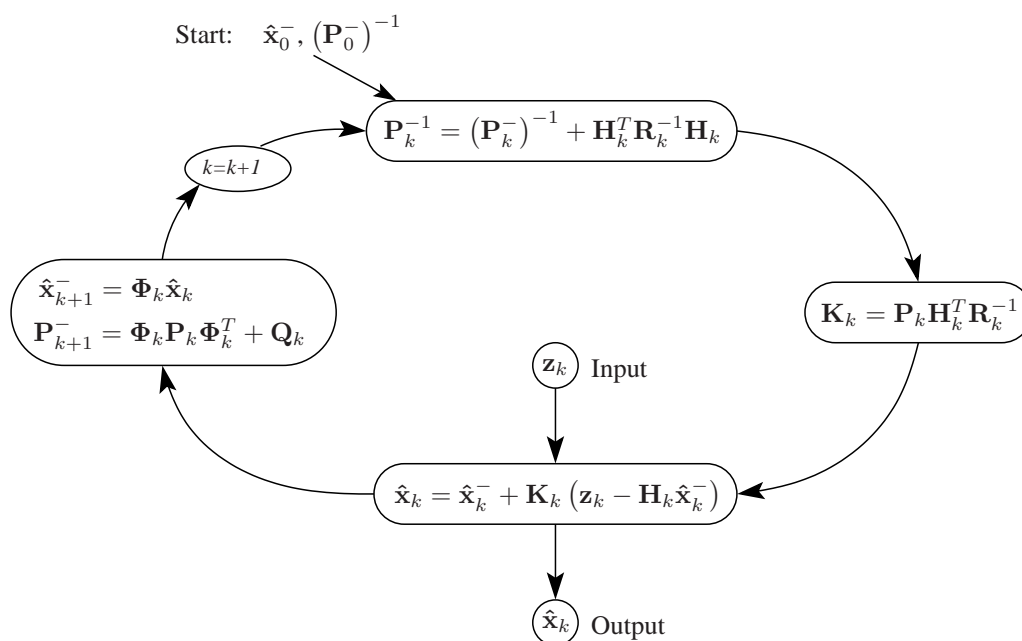


Abbildung 3.3: Alternativer Kalman-Filteralgorithmus

Anwendungsgebiete des Alternativen Kalman-Filters

1. Es ist sinnvoll den Kalman-Filter in seiner alternativen Schreibweise zu verwenden, falls keine Vorahnung über die zu suchende Zustandsgröße \mathbf{x}_k existiert. Stochastisch ausgedrückt bedeutet dies, daß die Streuungen $\sigma_{e_{0i}}^2$ des Schätzfehlers \mathbf{e}_0^- und damit auch die Einträge auf der Hauptdiagonalen der zugehörigen Kovarianzmatrix \mathbf{P}_0^- unendlich sind.

In der Gleichung (3.42) des regulären Filters zur Berechnung der Kalman-Verstärkungsmatrix führt diese Startbedingung zu dem unbestimmten Ausdruck ∞ . Die alternative Schreibweise des Kalman-Filters verwendet hingegen als Initialbedingung die Inverse der Fehlerkovarianzmatrix $(\mathbf{P}_0^-)^{-1}$, so daß keine unbestimmten Ausdrücke für den zuvor genannten Fall auftreten, da gilt

$$\mathbf{P}_0^- = \infty \quad \rightarrow \quad (\mathbf{P}_0^-)^{-1} = \mathbf{0}. \quad (3.51)$$

2. Beim Kalman-Filter in seiner regulären Form muß zur Berechnung der Kalman-Verstärkungsmatrix die Matrix $(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)$ invertiert werden. Sie beinhaltet m Zeilen- und Spalteneinträge, wobei m die Anzahl der Meßwerte z_j ist.

Für das alternative Kalman-Filter hingegen muß während des Iterationsverlaufes die Inverse der Matrix \mathbf{P}_k mit der Mächtigkeit $n \times n$ gebildet werden, mit n als Anzahl der zu schätzenden Zustandsgrößen x_i .

Bei der numerischen Invertierung von Matrizen wächst der Rechenaufwand bei linear ansteigender Mächtigkeit der Matrix kubisch. Daraus ergibt sich, daß für die Anwendungsfälle, in denen die Anzahl der Meßwerte groß gegenüber der Zustandsgrößen ist, also $n \ll m$ gilt, die alternative Form des Kalman-Filters besser geeignet ist, da sich in dieser der numerische Aufwand für die Invertierung der Matrizen erheblich reduziert.

3.2.4 Nichtlineare Kalman-Filterung

Da viele Modelle dynamischer Prozesse sowohl in ihren Systemgleichungen als auch in den Abbildungen der Meßgrößen nichtlineare Beziehungen beinhalten, das lineare Kalman-Filter diese aber nicht berücksichtigt, ist es notwendig, die Problemstellung diesbezüglich anzupassen. Nachfolgend werden zwei Arten der Linearisierung des Kalman-Filters nach [8] kurz vorgestellt: das linearisierte und das erweiterte Kalman-Filter. Eine ausführliche Behandlung dieser beiden Formen des Kalman-Filters findet in [25] statt.

Linearisiertes Kalman-Filter

Der zu schätzende dynamische Prozeß und die zugehörige Meßabbildung seien in folgender zeitkontinuierlicher Form darstellbar

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}_d, t) + \mathbf{w}(t), \quad (3.52)$$

$$\mathbf{z} = \mathbf{h}(\mathbf{x}, t) + \mathbf{v}(t). \quad (3.53)$$

Hierbei sind \mathbf{f} und \mathbf{h} bekannte Funktionen, die nichtlinear sein dürfen, \mathbf{u}_d ist ein deterministisches Eingangssignal. Die Größen \mathbf{w} und \mathbf{v} sind weiße Störprozesse des Signals und gegenseitig unkorreliert. Die Störprozesse werden als separate Terme angenommen.

Es wird nun ein beliebig angenäherter zeitlicher Verlauf der Zustandsgröße $\mathbf{x}^*(t)$ ermittelt, der nominale Trajektorie genannt wird und in Abbildung 3.4 dargestellt ist.

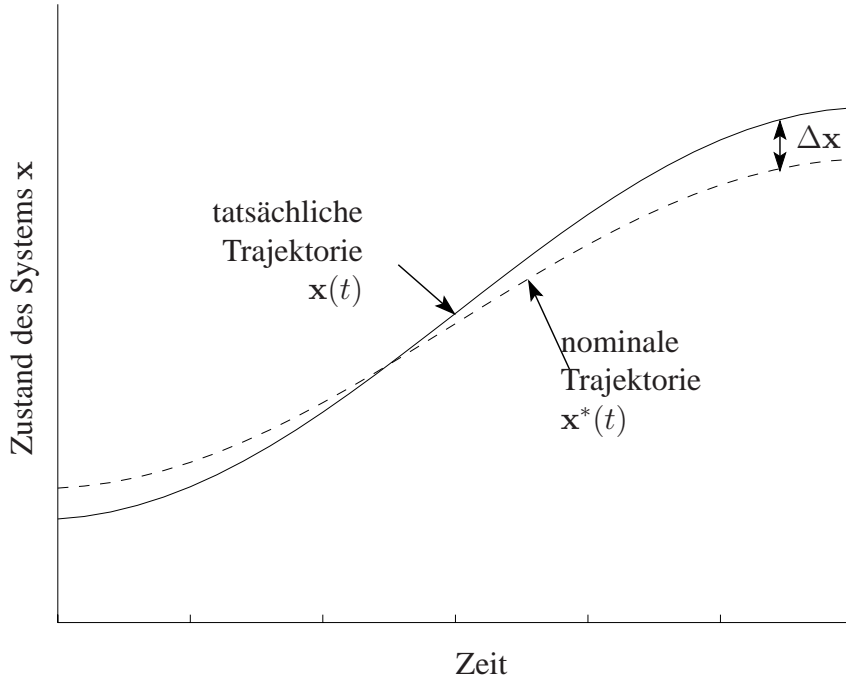


Abbildung 3.4: Tatsächliche und nominale Trajektorie des linearisierten Kalman-Filters

Der grundlegende Gedanke des linearisierten Kalman-Filters ist, den zeitlichen Verlauf der Zustandsgröße \mathbf{x} nicht absolut, sondern relativ zu der nominalen Trajektorie \mathbf{x}^* zu schätzen. Für die Betrachtung der Änderung von \mathbf{x} gegenüber \mathbf{x}^* kann eine Sensitivitätsanalyse erfolgen, mit der das Filter linearisiert wird.

Die tatsächlich existierende Trajektorie kann mit Hilfe von \mathbf{x}^* wie folgt geschrieben werden

$$\mathbf{x}(t) = \mathbf{x}^*(t) + \Delta \mathbf{x}(t). \quad (3.54)$$

Mit dieser Schreibweise werden die Gleichungen (3.52) und (3.53) zu

$$\dot{\mathbf{x}}^* + \Delta \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}^* + \Delta \mathbf{x}, \mathbf{u}_d, t) + \mathbf{w}(t), \quad (3.55)$$

$$\mathbf{z} = \mathbf{h}(\mathbf{x}^* + \Delta \mathbf{x}, t) + \mathbf{v}(t). \quad (3.56)$$

Werden die Funktionen \mathbf{f} und \mathbf{h} mit Taylor Reihen angenähert und nur die Terme erster Ordnung berücksichtigt, ergeben sich unter der Annahme, daß $\Delta \mathbf{x}$ klein ist

$$\dot{\mathbf{x}}^* + \Delta \dot{\mathbf{x}} \approx \mathbf{f}(\mathbf{x}^*, \mathbf{u}_d, t) + \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right]_{\mathbf{x}=\mathbf{x}^*} \cdot \Delta \mathbf{x} + \mathbf{w}(t), \quad (3.57)$$

$$\mathbf{z} \approx \mathbf{h}(\mathbf{x}^*, t) + \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right]_{\mathbf{x}=\mathbf{x}^*} \cdot \Delta \mathbf{x} + \mathbf{v}(t). \quad (3.58)$$

Hierbei symbolisieren $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ und $\frac{\partial \mathbf{h}}{\partial \mathbf{x}}$ die Jacobi-Matrizen

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}; \quad \frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (3.59)$$

Üblicherweise wird die nominale Trajektorie so bestimmt, daß die anschließende deterministische Differentialgleichung erfüllt ist

$$\dot{\mathbf{x}}^* = \mathbf{f}(\mathbf{x}^*, \mathbf{u}_d, t). \quad (3.60)$$

Die Verwendung dieser Formulierung in Gleichung (3.57) führt zum linearisierten Prozeßmodell

$$\Delta \dot{\mathbf{x}} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right]_{\mathbf{x}=\mathbf{x}^*} \cdot \Delta \mathbf{x} + \mathbf{w}(t), \quad (3.61)$$

$$\mathbf{z} - \mathbf{h}(\mathbf{x}^*, t) = \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right]_{\mathbf{x}=\mathbf{x}^*} \cdot \Delta \mathbf{x} + \mathbf{v}(t). \quad (3.62)$$

Das linearisierte Kalman-Filter schätzt den Abstand $\Delta \mathbf{x}$ der Zustandsgröße \mathbf{x} von der nominalen Trajektorie \mathbf{x}^* und nicht die absoluten Werte. Ebenso entsprechen die Meßdaten des linearisierten Modells nur der Differenz der tatsächlich gemessenen Werte \mathbf{z} und der durch die nominale Trajektorie vorausgesagten $\mathbf{h}(\mathbf{x}^*, t)$. Und auch die Jacobi-Matrizen aus Gleichung (3.59) beinhalten die partiellen Ableitungen, die an der nominalen Trajektorie ausgewertet wurden.

Erweitertes Kalman-Filter

Das erweiterte Kalman-Filter funktioniert ähnlich dem linearisierten mit dem Unterschied, daß die Linearisierung hier nicht anhand einer zuvor ermittelten nominalen Trajektorie stattfindet. Vielmehr werden die partiellen Ableitungen aus Gleichung (3.59) an einer Trajektorie ausgewertet, die mit den Schätzwerten $\hat{\mathbf{x}}_k^-$ des Filters aktualisiert wird. Da die Schätzwerte wiederum von den Meßwerten \mathbf{z}_k abhängen, wird deutlich, daß die Filter-Sequenz von einem bestimmten Durchlauf des Experimentes abhängig ist. Die Filter-Abfolge ist daher im Unterschied zum linearen Kalman-Filter nicht durch die Annahmen des Prozeßmodells a priori vorgegeben.

In jedem betrachteten Zeitschritt wird eine Sensitivitätsanalyse auf Grundlage der statistisch verbesserten Trajektorie \mathbf{x}^* durchgeführt. Annahmen für den geschätzten Verlauf der Zustandsgröße können so während des Algorithmus korrigiert werden, falls sie die Optimierung behindern.

Die geschätzte Trajektorie wird, wie in Abbildung 3.5 dargestellt, in jedem Schritt statistisch verbessert, was allerdings keine reale Verbesserung der Schätzung bedeuten muß. Es kann sogar im Gegenteil zu einer Verschlechterung kommen. Das erweiterte Kalman-Filter birgt das Risiko einer schlechteren Schätzung gegenüber dem linearen, da die Schätzfehler durch ungünstig angenommene Trajektorien möglicherweise noch verstärkt werden.

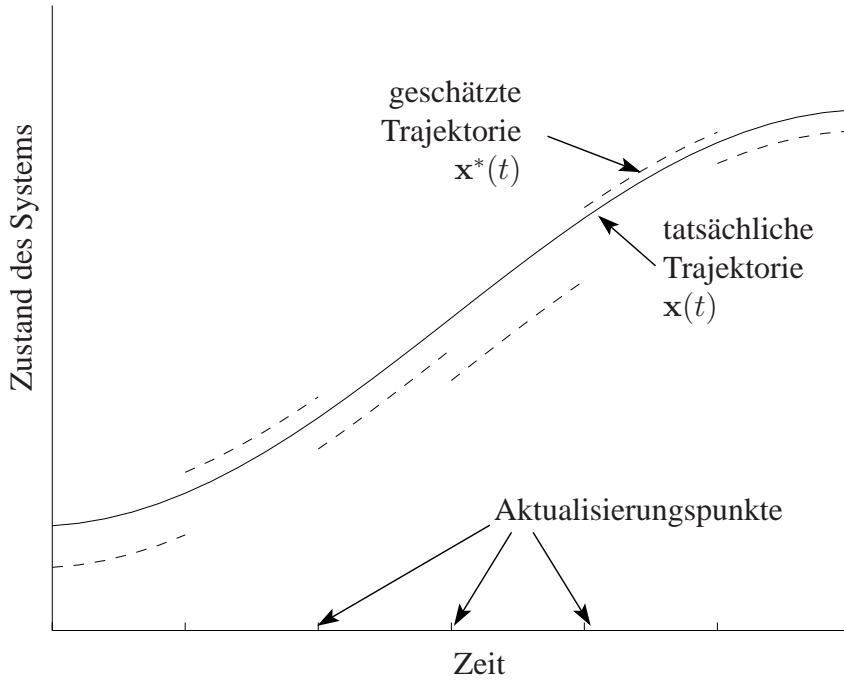


Abbildung 3.5: Tatsächliche und geschätzte Trajektorie des erweiterten Kalman-Filters

Das Symbol \mathbf{x}^* stellt nun die in jedem diskreten Zeitschritt aktualisierte, geschätzte Trajektorie dar. Die Matrizen aus Gleichung (3.59) seien für das erweiterte Kalman-Filter

$$\mathbf{F}_k = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right]_{\mathbf{x}_k = \mathbf{x}_k^*} \quad \text{und} \quad \mathbf{H}_k = \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right]_{\mathbf{x}_k = \mathbf{x}_k^*} \quad (3.63)$$

Beim linearisierten Kalman-Filter sind die Zustandsgrößen inkrementeller Art und nicht absolut. Es wird nur eine Aussage über ihre Änderung getroffen. Da es allerdings üblich ist, die absoluten Größen zu beobachten, wird hier gezeigt, wie die inkrementelle Darstellung in die absolute überführt werden kann.

Ausgehend von der linearisierten Meßgleichung (3.62)

$$\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^*) = \mathbf{H}_k \cdot \Delta \mathbf{x}_k + \mathbf{v}_k \quad (3.64)$$

zur Darstellung der inkrementellen, aktualisierten Schätzung zum Zeitpunkt t_k , wird $\Delta \hat{\mathbf{x}}_k$ berechnet zu

$$\Delta \hat{\mathbf{x}}_k = \Delta \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^*) - \mathbf{H}_k \Delta \hat{\mathbf{x}}_k^-). \quad (3.65)$$

Da die prognostizierte Schätzung der Meßgröße $\hat{\mathbf{z}}_k^-$ die Summe von $\mathbf{h}(\mathbf{x}_k^*)$ und $\mathbf{H}_k \Delta \hat{\mathbf{x}}_k^-$ ist, wird die Kalman-Verstärkungsmatrix \mathbf{K}_k also mit dem Residuum der Meßwerte $(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)$ multipliziert.

Aus Addition der Gleichung (3.65) und \mathbf{x}_k^* entsteht

$$\mathbf{x}_k^* + \Delta \hat{\mathbf{x}}_k = \mathbf{x}_k^* + \Delta \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k^-). \quad (3.66)$$

oder anders ausgedrückt nach Gleichung (3.54)

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k^-). \quad (3.67)$$

Gleichung (3.67) ist die bekannte Aktualisierungsgleichung der Schätzung, die die absoluten Werte der Zustandsgrößen verwendet. Sie besagt, daß die a priori Schätzung durch das mit der Kalman-Verstärkungsmatrix gewichtete Residuum der Meßwerte korrigiert wird.

3.3 Genetischer Algorithmus

Genetische Algorithmen zählen zu den heuristischen Suchverfahren. Eine ausführlichere Einleitung zu den Funktionsweisen genetischer Algorithmen und Evolutionsstrategien kann in [38] und [3] nachgelesen werden. Zu Beginn des Algorithmus wird eine Population erstellt. Diese Population ihrerseits besteht aus n Individuen, welche aus Chromosomen zusammengesetzt sind. Abbildung 3.6 zeigt einen exemplarischen Aufbau eines Individuums mit vier Genen. Die-

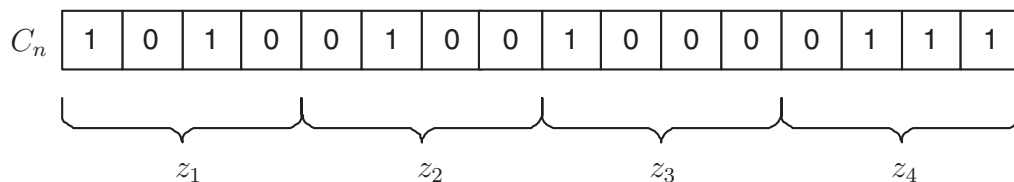


Abbildung 3.6: Darstellung eines binär kodierten Chromosoms mit vier Genen

sen Genen werden hier die gesuchten, zu optimierenden Parameter z_i zugeordnet. Bei der Schadenserkennung sind dies die Parameter der geometrischen Größen der Schädigung. Innerhalb des genetischen Algorithmus werden diese reellwertigen Parameter durch eine Transformation jeweils in eine binärwertige Zahl umgewandelt. Dabei wird der Wertebereich der Parameter zu Beginn des Algorithmus sinnvoll eingeschränkt. So sollte sich der Bereich für die Mittelpunktkoordinaten innerhalb der Geometrie befinden, da keine Fehlstellen außerhalb der Geometrie liegen können.

Der genetische Algorithmus versucht, den sogenannten Fitneßwert, vgl. (3.1), zu optimieren. Der Fitneßwert hängt direkt mit den zu suchenden Parametern zusammen. Je nach Aufgabenstellung muß versucht werden, eine geeignete Formulierung für den Fitneßwert mit den zu suchenden Größen als Parameter zu erstellen. Es muß unterschieden werden, ob der Fitneßwert maximiert oder minimiert werden soll. Bei einem Minimierungsproblem soll ein möglichst kleiner Wert erzielt werden.

In [38] werden viele Beispiele vorgestellt, die durch den genetischen Algorithmus optimiert werden.

Im ersten Schritt bei Anwendung des genetischen Algorithmus wird eine Anfangspopulation aus Individuen erstellt. Die Individuen bestehen aus zufällig angenommenen Größen der zu suchenden Parameter. Daraufhin wird von allen Individuen der Anfangspopulation deren jeweilige Fitneß bestimmt. Mit Hilfe dieses Fitneßwertes werden die Individuen der gesamten Population

sortiert. Das Individuum mit der besten Fitneß steht an erster Stelle, gefolgt von demjenigen mit dem zweitbesten Fitneßwert. Individuen mit schlechten Fitneßwerten bilden den Schluß der sortierten Reihe. Nichtzulässigen Parameterkombinationen sollte ein sehr schlechter Fitneßwert zugewiesen werden. Bei einem Minimierungsproblem entspricht dies einem sehr großen Zahlenwert. Die folgenden Schritte haben einen iterativen Charakter. Aus dieser Liste von Individuen werden die n_r besten ausgewählt, die die Elterngeneration darstellen. Auf diese Individuen werden nun genetische Operatoren angewandt. Im einzelnen sind dies:

- Chromosomentausch

Für den Chromosomentausch werden jeweils zwei Individuen aus der Menge der Eltern ausgewählt und mit der Wahrscheinlichkeit p_{CT} wird das m_{CT} -te Chromosom ausgetauscht.

- Crossing-Over

Beim Crossing-Over wird mit zwei Individuen der Elterngeneration analog vorgegangen. Der Operator wird auf jedes Chromosom mit der Wahrscheinlichkeit p_{CO} angewandt. Dabei wird an der zufällig bestimmten Position im Chromosom der Chromosomenstrang geöffnet, die rechten Enden des Stranges zwischen den beiden ausgewählten Individuen werden ausgetauscht und dann zusammengefügt. Unterschieden wird zwischen einem einfachen Crossover und einem Zwei-Punkt-Crossover. Der Unterschied liegt in der Anzahl der Trennungen innerhalb des Individuums. In Abbildung 3.7 wird das Individuum nur an einer durch den Zufall bestimmten Position aufgetrennt und neu zusammengesetzt, während Abbildung 3.8 den Vorgang des Zwei-Punkt-Crossover abbildet. Hier wird an zwei beliebigen Positionen innerhalb des Individuums die binäre Kodierung aufgespalten und danach neu über Kreuz zusammengefügt.

- Mutation

Der Mutationsoperator wird, im Gegensatz zu den zuvor beschriebenen Operatoren, auf einzelne Individuen der Elterngeneration angewandt. Der Mutationsoperator wird auf jedes Chromosom mit der Wahrscheinlichkeit p_M angewandt. Dabei wird an der k -ten Stelle im Chromosom das betreffende Bit negiert; k ist jeweils zufällig ermittelt. Anschaulich ist dieser Vorgang in Abbildung 3.9 dargestellt.

Bevor diese Operatoren alle Individuen der Elterngeneration beeinflussen können, wird die Liste der Elterngeneration zufällig umsortiert, damit beim Chromosomentausch und beim Crossing-Over möglichst neue Individuen entstehen. Sollten doch gleiche Individuen (Clone) entstehen, werden diese aus der Liste der neuen Generation aussortiert. Nun wird auf diese Menge der Nachkommenschaft die Fitneßfunktion angewandt. Anschließend wird diese Menge von Individuen der Menge, welche zu Beginn des Iterationsschrittes vorliegt, hinzugefügt. Die erweiterte Population wird bezüglich der Fitneßwerte wieder geordnet. Aus dieser neu erstellten Liste werden wieder die n besten Individuen ausgewählt, welche die neue Generation bilden. Damit ist ein Iterationsschritt beendet. Die Iteration des genetischen Algorithmus ist dann abgeschlossen, wenn der beste Fitneßwert einen vorgegebenen Wert erreicht hat, oder die maximale Anzahl an Iterationen vollzogen ist. Anzumerken ist, daß eine Vielzahl unterschiedlicher Varianten eines

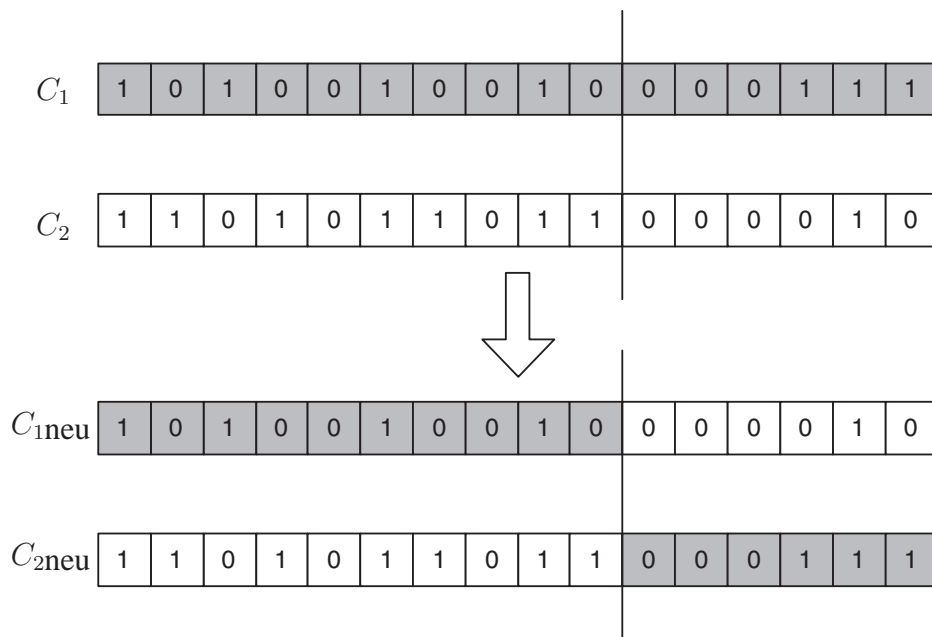


Abbildung 3.7: Darstellung eines Ein-Punkt-Crossover

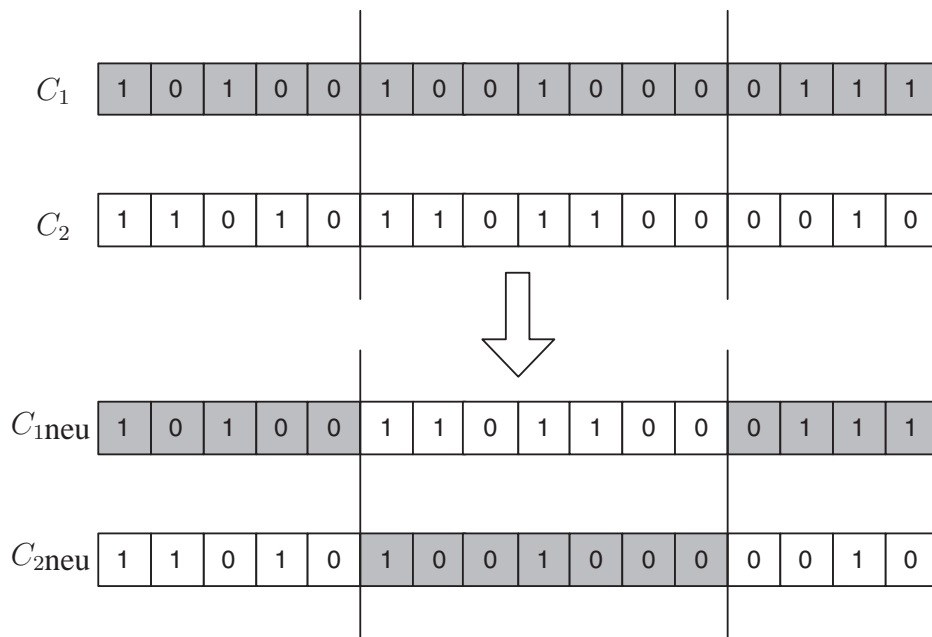


Abbildung 3.8: Darstellung eines Zwei-Punkt-Crossover

genetischen Algorithmus zur Verfügung stehen. Entsprechende Eigenschaften dieser Algorithmen können in [38], [4] und [5] nachgelesen werden.

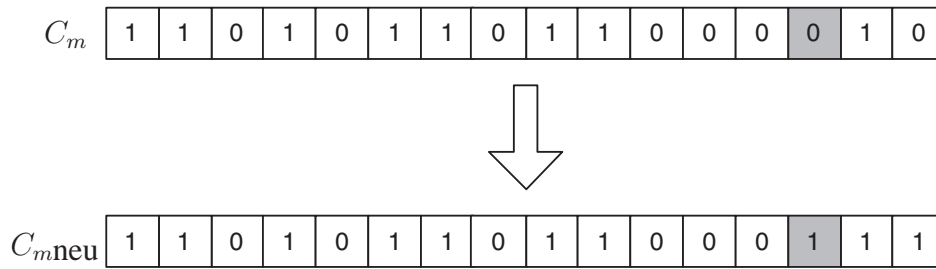


Abbildung 3.9: Mutation innerhalb eines Individuums

3.4 Neuronales Netz

Neuronale Netze sind informationsverarbeitende Systeme, die aus einer großen Anzahl einfacher Einheiten, den sogenannten Neuronen, bestehen. Ein neuronales Netzwerk entsteht durch die gewichteten Verbindungen zwischen diesen Neuronen. Die Sendung von Informationen zwischen den Neuronen erfolgt durch diese Verbindungen in Form von Aktivierungspotentialen. Künstliche neuronale Netze ähneln den biologischen neuronalen Netzen. Die Informationsverarbeitung findet in sehr vielen Nervenzellen gleichzeitig statt. Den Grad ihrer Erregung geben sie über Nervenfasern an andere Nervenzellen weiter. Erst durch ihre massiv parallele Arbeitsweise erhalten neuronale Netze ihre Problemlösefähigkeit. Ein wesentliches Element eines neuronalen Netzes ist die Lernfähigkeit. Aus Trainingsbeispielen sind sie in der Lage, Klassifikationsprobleme und Optimierungsprobleme selbständig zu lösen, ohne dabei neu programmiert werden zu müssen. Die Struktur des Netzes versetzt sie auch in die Lage, komplexe und zunächst verborgene Zusammenhänge zu erkennen. Eine ausführlichere Beschreibung der Arbeitsweise neuronaler Netze kann [52] entnommen werden.

Zur Lösung der Aufgabenstellung der Schädigungsidentifizierung werden nur vollständig verbundene feed-forward Netzarchitekturen auf Basis des Backpropagation-Algorithmus verwendet. Bei einem vollständig verbundenen feedforward-Netz ist jedes Neuron einer Schicht mit dem Neuron in der nächst höher gelegenen Schicht verbunden. Ein „Backpropagationnetz“ nach Abbildung 3.10 besteht aus einer Eingabeschicht, einer Ausgabeschicht und mindestens einer verborgenen Schicht. In der Eingabeschicht befinden sich die R Eingangsneuronen a^0 , denen die Daten von außen mit den Lerndaten \mathbf{p} zugeführt werden. Die Anzahl der Zwischenschichten und die Anzahl der Neuronen s je Zwischenschicht ist variabel. In den meisten Anwendungen genügt eine verborgene Zwischenschicht. In der Ausgabeschicht befinden sich die Ausgangsneuronen, die das prognostizierte Ergebnis \mathbf{t} liefern sollen. Um auch stark nichtlineare Zusammenhänge zu erkennen, wird dem Ausgang eines einzelnen Neurons nach Abbildung 3.11 eine Aktivierungsfunktion f vorgeschaltet. Als Aktivierungsfunktion stehen mehrere Funktionen zur Auswahl. Zur Schädigungsidentifizierung wird der Tangens Hyperbolicus gewählt, da dieser auch negative Ausgänge der Neuronen berücksichtigen kann. In Abbildung 3.11 wird ein einzelnes Neuron in der ersten verborgenen Schicht betrachtet. Der Ausgang a_1^2 des Neurons

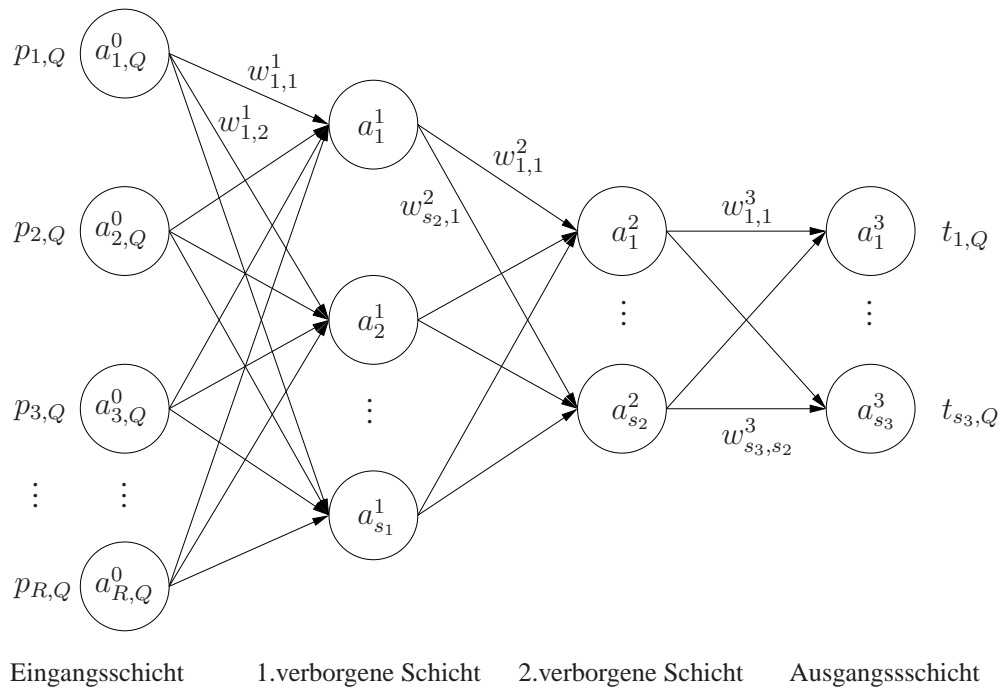


Abbildung 3.10: Backpropagationnetzwerk mit zwei verborgenen Schichten

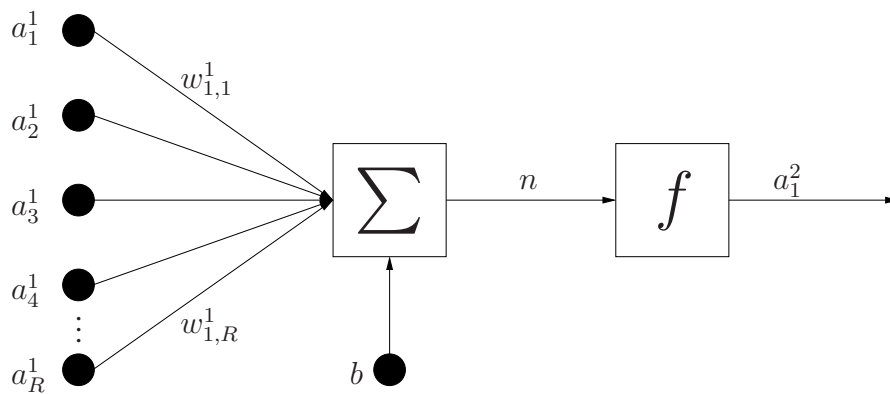


Abbildung 3.11: Beispielhafter Aufbau eines einzelnen Neurons in der ersten Schicht

berechnet sich aus Gleichung

$$a_1^2 = f^1(n), \quad (3.68)$$

in der n aus der Summe der einzelnen Ausgänge der R Neuronen aus der Eingabeschicht besteht

$$n = \sum_{l=1}^{l=R} a_l^1 \cdot w_{1,l}^1 + b. \quad (3.69)$$

b kennzeichnet den Bias-Wert, der einen zusätzlichen unbekannten Parameter des betrachteten Neurons darstellt. Darauf aufbauend kann das Übertragungsverhalten der ersten Schicht nach

$$\mathbf{a}_1^2 = f^1(\mathbf{W}^1 \mathbf{a}^1) \quad (3.70)$$

berechnet werden, in der \mathbf{W}^1 ein Vektor darstellt, in der alle einzelnen Werte der Wichtungen zwischen Eingabe- und ersten Zwischenschicht enthalten sind. In [11] können weitere Einzelheiten der Anpassung des neuronalen Netzes an das Problem der Fehlstellenidentifizierung entnommen werden. Eine ausführliche Beschreibung der unterschiedlichen Netzarchitekturen ist in [35] beschrieben.

Das Backpropagation-Verfahren ist ein iteratives Verfahren. Es ermittelt eine Konfiguration der Gewichte im Netz, bei der die Fehlerfunktion über alle Trainingsmuster minimal ist. Als Fehlerfunktion wird sehr häufig der mittlere quadratische Fehler (MSE) benutzt, der die reale Ausgabe mit der erwarteten Ausgabe des Netzes vergleicht [52]. Künstliche neuronale Netze auf der Basis von Backpropagation werden zum jetzigen Zeitpunkt für praktische Anwendungen am häufigsten verwendet. Der Lernvorgang des neuronalen Netzes erfolgt in einem Batchbetrieb. Die Trainingsdaten in Form von den errechneten Größen liegen alle bei Trainingsbeginn des Netzes vor. Zusätzlich werden die Ausgangsgrößen einem neuronalen Netz zur Verfügung gestellt. Dieser Lernvorgang wird auch als „Überwachtes Lernen“ bezeichnet. Zu jedem Trainingsbeispiel wird dem Netz die richtige Ausgangsgröße gegenübergestellt, so daß der entstehende Fehler nach der Fehlerfunktion berechnet werden kann. Der Lernvorgang eines neuronalen Netzes kann in drei Schritten unterteilt werden. Zuvor muß allerdings die für den Lernvorgang des Netzes notwendige Lernmenge $\{\mathbf{p}_1, \mathbf{t}_1\}, \{\mathbf{p}_2, \mathbf{t}_2\}, \dots, \{\mathbf{p}_Q, \mathbf{t}_Q\}$, mit Q Lernbeispielen erstellt werden. Hier gilt es geeignete Lernbeispiele auszusuchen, die den Erfolg des Trainings garantieren. Es ist immer hilfreich, sich die konkrete Anwendung genau anzuschauen und eine geeignete Vorauswahl der Lernmenge zu treffen. Weitere Einzelheiten zur Erstellung der Lernmenge kann unter anderem in [29] und [35] nachgelesen werden. Desweiteren ist eine Testmenge \mathbf{p}^T, \mathbf{z} zu berechnen, durch die die Leistungsfähigkeit des Netzes eingeschätzt werden kann. Damit ergibt sich der erste Schritt zum Aufbau des Netzes

- Vorwärtsschritt

Mit dem Lerndatensatz bestehend aus Eingangsgröße \mathbf{p} wird die Netzausgabe bestimmt. Nach Abbildung 3.10 ergibt sich hierfür als Ausgang \mathbf{a}^3 des Netzes

$$\mathbf{a}^3 = f^3(\mathbf{W}^3 f^2(\mathbf{W}^2 f^1(\mathbf{W}^1 + \mathbf{b}^1) + \mathbf{b}^2) + \mathbf{b}^3). \quad (3.71)$$

- Fehlerberechnung

Die Netzausgabe wird mit den Sollgrößen verglichen. Hierzu dient der mittlere quadratische Fehler berechnet nach

$$F(\mathbf{x}) = (\mathbf{t}(k) - \mathbf{a}(k))^T (\mathbf{t}(k) - \mathbf{a}(k)). \quad (3.72)$$

In Gleichung (3.72) stellt \mathbf{x} einen Vektor bestehend aus den unbekannten Wichtungen und der Bias-Werte mit

$$\mathbf{x} = [\mathbf{W}^1; \mathbf{b}^1; \mathbf{W}^2; \mathbf{b}^2; \dots; \mathbf{W}^M; \mathbf{b}^M]^T \quad (3.73)$$

bei insgesamt M verdeckten Schichten dar.

- Rückwärtsschritt (Backpropagationsschritt)

Die Gewichte des neuronalen Netzes werden ausgehend von der Ausgabeschicht bis hin zur Eingangsschicht einer Lernregel entsprechend geändert. Hierzu wird sich einer Optimierungsstrategie bedient. Der klassische Backpropagationalgorithmus verwendet ein einfaches Gradientenabstiegsverfahren, in dem die Gewichte in Richtung des steilsten Gradienten der Fehlerfunktion geändert werden [52]. Die notwendigen Gleichungen hierzu ergeben sich zu

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \alpha \frac{\partial F}{\partial w_{i,j}^m}, \quad (3.74)$$

$$b_i^m(k+1) = b_i^m(k) - \alpha \frac{\partial F}{\partial b_i^m}, \quad (3.75)$$

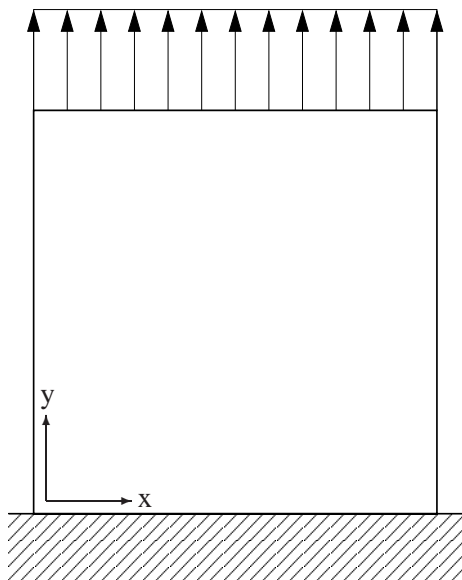
in denen α die Lernrate des Optimierungsverfahrens darstellt und die Schrittweite bestimmt, mit dem die zu optimierenden Parameter verbessert werden. Die partiellen Ableitungen in den Gleichungen (3.74) und (3.75) müssen berechnet werden [52].

Der Trainingsvorgang des Netzes gilt als beendet, wenn eine Abbruchbedingung erfüllt ist. Zum einen kann eine vorgegebene Anzahl an Iterationen k zum Abbruch des Lernvorganges führen. Eine andere Abbruchbedingung ist die Unterschreitung einer vorgegebenen Fehlergrenze für $F(\mathbf{x})$.

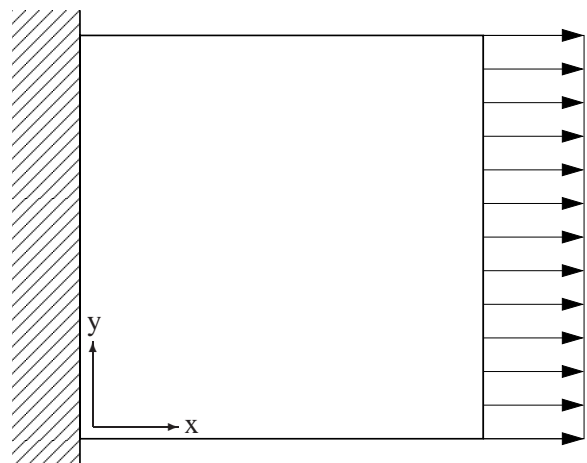
Eine Besonderheit bei Anwendung neuronaler Netze ist das sogenannte „Overfitting“. Das neuronale Netz wird zu lange trainiert und gibt den Testdatensatz nicht genau genug wieder, obwohl das Training perfekt beendet wird. Um dieses Resultat zu meiden empfiehlt [29] folgende Regel zu beachten. Die Anzahl der unbekannten Wichtungen und Bias-Werte sollen die Anzahl der Lerndatensätze Q nicht überschreiten.

Der Vorteil des neuronalen Netzes gegenüber anderen Optimierungsstrategien zur Fehlstellenidentifizierung ist die Anwendung. Ein einmal trainiertes Netz kann für alle Testbeispiele angewandt werden und muß für diesen Anwendungsfall nicht neu angelern werden.

4 Loch- und Rißerkennung unter statischer Belastung



(a) Lastfall 1



(b) Lastfall 2

Abbildung 4.1: Die verwendeten Lastfälle bei einer beispielhaft durch einen Riß geschädigten Scheibe

Die vorgestellten Optimierungsstrategien werden in dieser Studie zur Schadenserkennung in zweidimensionalen und dreidimensionalen Geometrien angewandt. Die iterativen Optimierungsstrategien müssen mit der Randelementsimulation gekoppelt werden. Dabei wird eine Parametrisierung der Randelementmethode durchgeführt. Da auch mehrere Fehlstellen innerhalb einer Geometrie auftreten können, besteht die Möglichkeit, daß sich Fehlstellen zum Teil überlappen. Dieses Überschneiden möglicher Fehlstellen wird kontrolliert und unterbunden, da die neu entstandene Form der Schädigung nicht in der Parametrisierung der Randelementsimulation berücksichtigt wird. Bevor die Identifikation durchgeführt werden kann, muß das „Experiment“

berechnet werden. Dabei wird eine Randelements simulation mit den zu suchenden Schädigungsparametern durchgeführt, und die Verformungsgrößen des Scheibenrandes werden gespeichert in u_0 . Da keine Laborexperimente vorliegen, werden für die Rißerkennung ausschließlich numerische Simulationen mit der Randelementmethode durchgeführt. Bis auf das neuronale Netz benötigen alle hier getesteten Optimierungsstrategien die Ergebnisse des Experimentes zum Erkennen der Fehlstelle während der Optimierung. Einzig das neuronale Netz wird mit Simulationsdaten trainiert und erst durch Auswertung des Testdatensatzes wird die Leistungsfähigkeit des Netzes überprüft. Eine korrekte Aussage über die Prognosefähigkeit des trainierten Netzes kann nur der Laborversuch wiedergeben. Die Zielgrößen der iterativen Optimierungsstrategien werden in den Abbildungen der Schädigungsparameter mit einer durchgezogenen Linie dargestellt. Im Folgenden werden die Materialdaten der elastostatischen Berechnungen mit einem Elastizitätsmodul von $E = 1000$ und einer Querkontraktionszahl von $\nu = 0,2$ festgelegt. Es werden die Auswirkungen mehrerer Lastfälle nach Abbildung 4.1 auf die Identifizierung untersucht.

4.1 Identifikation mit SQP

Gegenüber der Simulation mit dem genetischen Algorithmus erweist es sich bei Verwendung des SQP-Verfahrens als vorteilhaft, eine logarithmierte Darstellung des Fehlers

$$E(\mathbf{z}) = \ln \left(\sum_j^{2n} (u_j(\mathbf{z}) - u_{0j}(\bar{\mathbf{z}}))^2 + 10^{-6} \right) \quad (4.1)$$

zu betrachten. Durch diese Operation kann der Optimierungsvorgang verbessert werden. Zusätzlich wird ein sehr kleiner Wert zum Fehlerquadrat hinzuaddiert, um sehr große Funktionswerte im Fehlerverlauf zu vermeiden. Zudem ist der Funktionswert für $\ln 0$ nicht definiert. Dieser Wert tritt dann auf, wenn die Identifikation korrekt abgeschlossen ist. Der Fehlerwert hat nun den Wert $\ln 10^{-6} = -13,816$. Dieser untere Grenzwert ist in den Abbildungen des Fehlerverlaufes als waagerechte Linie eingezeichnet. In den Abbildungen, in denen die zu optimierenden Parameter den durchgeführten Iterationen gegenübergestellt sind, kennzeichnen waagerecht verlaufende, durchgezogene Linien die zu suchenden Zielgrößen.

Identifikation von elliptischen Fehlstellen

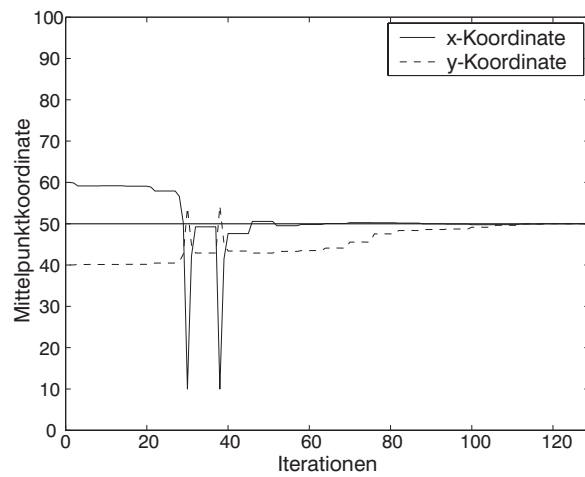
Die SQP-Methode wird zunächst mit einem Lastfall nach Abbildung 4.1a in einer Scheibe getestet. Die ersten Untersuchungen beschränken sich auf die Identifikation der Mittelpunktkoordinaten der jeweiligen Schädigung. Die Radien bzw. der Rißwinkel und die Rißlänge werden vorgegeben. Die Untersuchungen ergeben, daß die Mittelpunktkoordinaten einer elliptischen Schädigung oder eines Risses gut erkannt werden. Testrechnungen ergaben, daß der Startvektor der Optimierung, der die erste Annahme der zu suchenden Größen beinhaltet, keinen Einfluß auf die Identifikation hat. In diesen Versuchen fällt schon auf, daß viele Iterationsschritte für die Lokalisierung der Schädigung notwendig sind.

Bei Erweiterung der Problemstellung um die zwei Radien der elliptischen Fehlstelle auf vier Kenngrößen verläuft deren Prognose nicht zufriedenstellend. Die Optimierung ist nur erfolgreich, wenn der Startvektor aus Größen besteht, die eine Ellipse beschreiben, die bereits in der Nähe der zu suchenden Fehlstelle liegt. In Abbildung 4.2d ist die Übereinstimmung der gesuchten mit der identifizierten Ellipse deutlich zu erkennen. Der Fehlerverlauf, erkennbar in Abbildung 4.2c, konvergiert gut gegen den Grenzwert. Die Sprünge im Optimierungsverlauf resultieren vermutlich aus der Verwendung des Differenzenquotienten als Näherung der 1. Ableitung der Fehlerfunktion nach den Schädigungsparametern in der SQP-Methode.

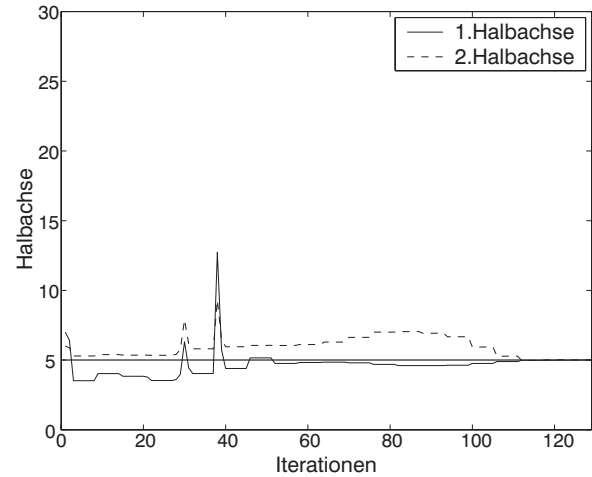
Die gleiche elliptische Schädigung soll bei Verwendung eines ungünstiger gewählten Startvektors erkannt werden. Der Verlauf der Fehlergröße, erkennbar in Abbildung 4.3c, ist deutlichen Schwankungen unterworfen. Das Ergebnis der Identifikation ist in Abbildung 4.3d ersichtlich. Die x-Koordinate ist nach Abbildung 4.3a sehr gut identifiziert worden, während die Länge einer Halbachse nicht erkannt wird.

Identifikation von rißförmigen Fehlstellen

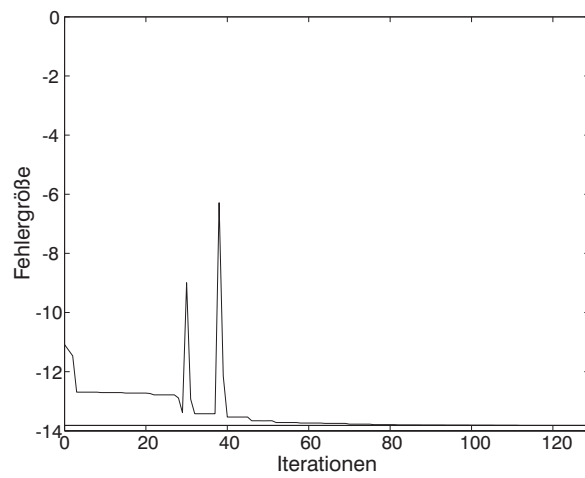
Der Optimierungsvorgang für die Identifikation von rißförmigen Fehlstellen unterscheidet sich nicht wesentlich von dem der elliptischen Fehlstelle. Die alleinige Prognose der Rißmittelpunktkoordinaten ist sehr gut. Bei Identifikationsversuchen von vier Parametern, den Mittelpunktkoordinaten und zusätzlich der Rißlänge und dem Rißwinkel, ist das Identifikationsergebnis auch hier abhängig von der Wahl des Startvektors. In Abbildung 4.4 ist beispielhaft die Identifikation zweier Risse dargestellt. Die Parameter des Startvektors haben eine geringe Abweichung von den zu identifizierenden Größen. So hat die Fehlergröße, abgebildet in Abbildung 4.4e, zu Beginn der Optimierung einen schon sehr kleinen Wert. Der Optimierungsvorgang an sich benötigt auf Grund der acht zu optimierenden Größen wesentlich mehr Iterationsschritte bis die Identifikation abgeschlossen ist als bei der Identifikation von nur einem Riß. Das Ergebnis der Optimierung kann erfolgreich abgeschlossen werden und kann Abbildung 4.4f entnommen werden.



(a) Mittelpunktkoordinaten



(b) Halbachsen



(c) Fehlerverlauf

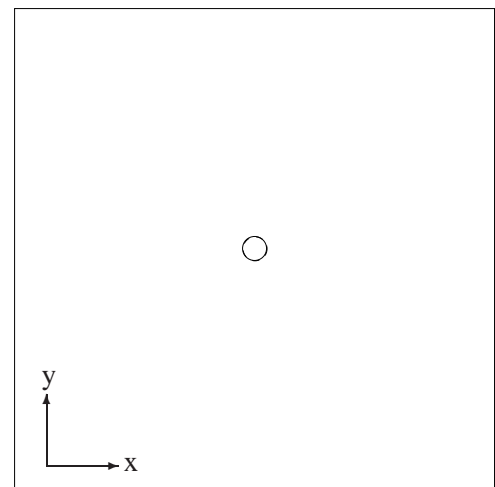
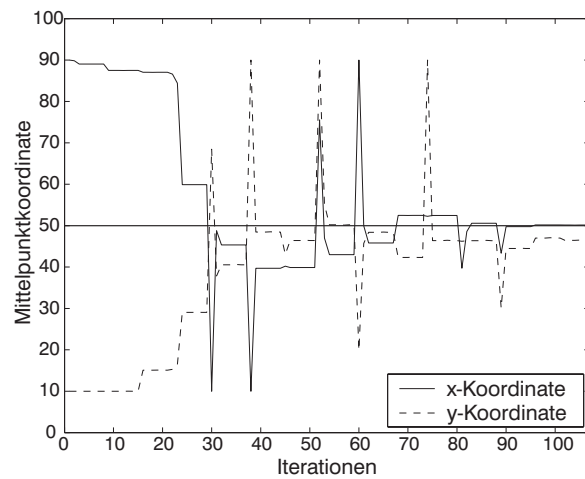
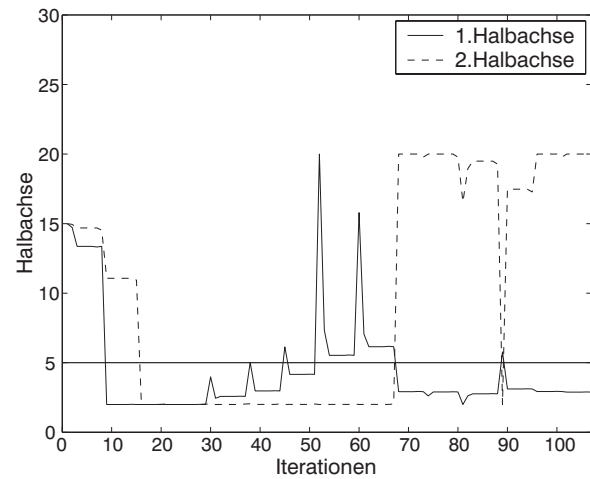

(d) Geometrie
--- Ergebnis — Vorgabe

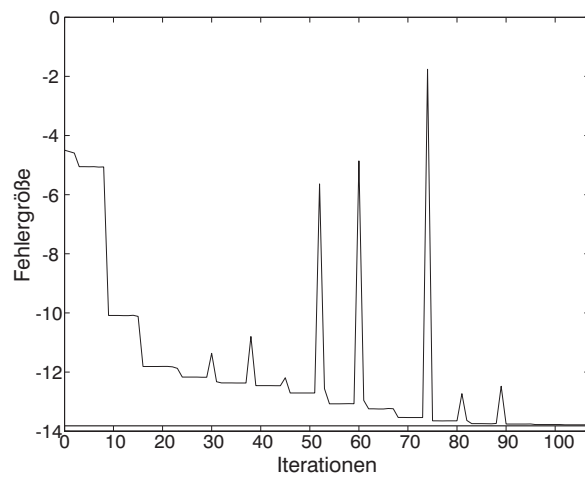
Abbildung 4.2: Identifikationsergebnis bei Verwendung des NAG SQP-Verfahrens mit einem von den zu suchenden Größen gering abweichenden Startwert von [60 40 6 7]



(a) Mittelpunktkoordinaten



(b) Halbachsen



(c) Fehlerverlauf

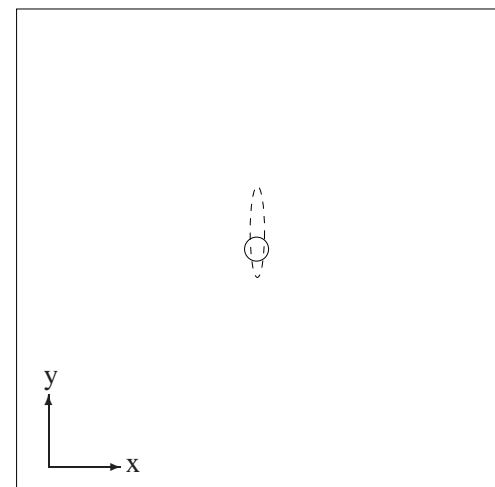
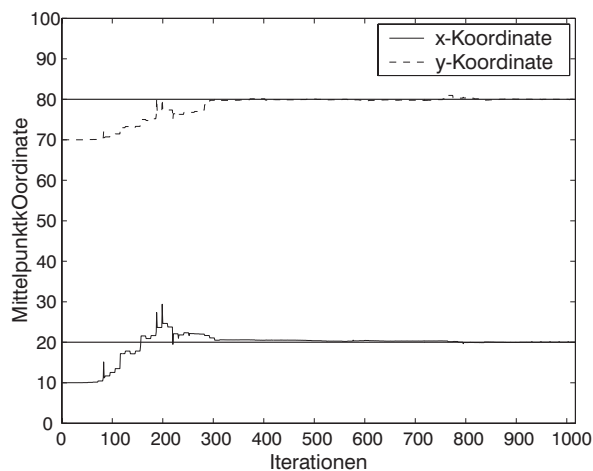
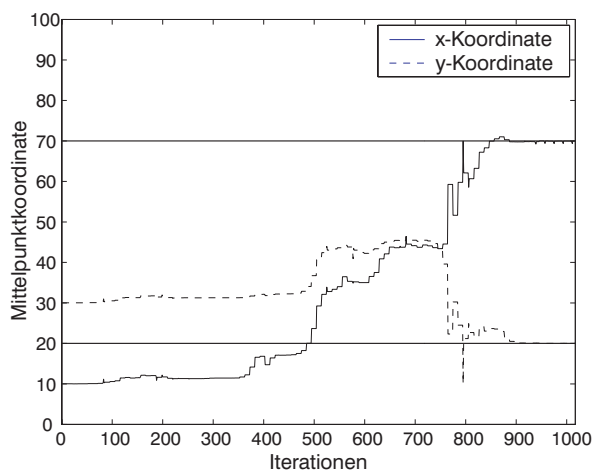
(d) Geometrie
- - - Ergebnis — Vorgabe

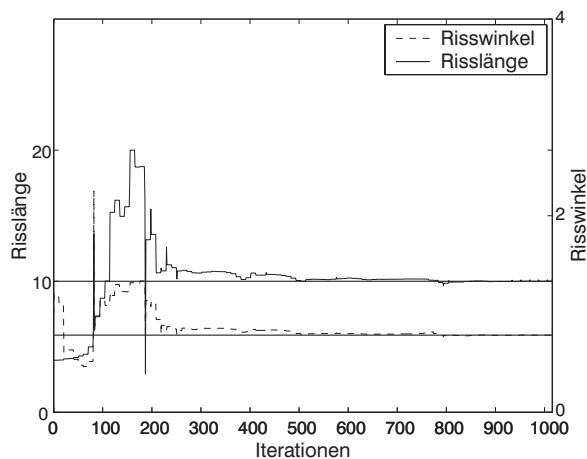
Abbildung 4.3: Identifikationsergebnis bei Verwendung des NAG SQP-Verfahrens mit einem von der Fehlstelle stark abweichenden Startwert von [90 10 15 15]



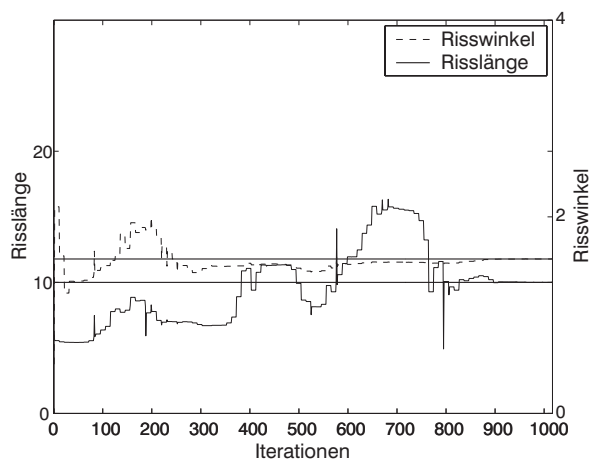
(a) Mittelpunktkoordinaten 1.Riß



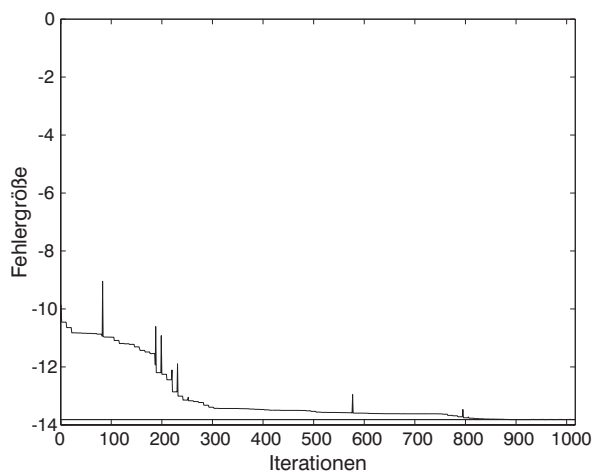
(b) Mittelpunktkoordinaten 2.Riß



(c) Rißlänge, Rißwinkel 1.Riß



(d) Rißlänge, Rißwinkel 2.Riß



(e) Fehlerverlauf

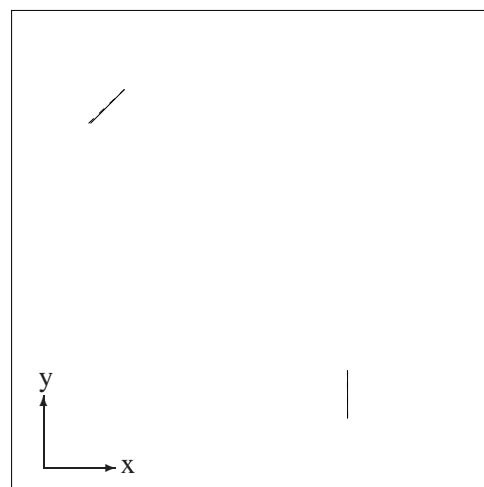

(f) Geometrie
---- Ergebnis ——— Vorgabe

Abbildung 4.4: Identifikation von zwei Rissen bei Verwendung des NAG SQP-Verfahrens mit den Startwerten für Riß 1: [10 70 4 1] und für Riß 2: [10 30 6 0,5]

4.2 Identifikation mit dem linearen Kalman-Filter

Um die Prognosefähigkeit des Kalman-Filters mit der SQP-Methode vergleichen zu können, werden die gleichen Untersuchungen wie im vorherigem Abschnitt durchgeführt. Es wird auch die gleiche logarithmierte Fehlerbeschreibung nach Gleichung (4.1) als Optimierungsfunktion verwendet wie im vorherigen Abschnitt.

Identifikation von elliptischen Fehlstellen

In Abbildung 4.5 ist der Optimierungsverlauf bei einer Identifikation eines mittig platzierten Loches dargestellt. Der Startwert wird gegenüber der SQP-Methode beibehalten, so daß die Ergebnisse vergleichbar sind. Bei Verwendung der gleichen Skalierung wie in den Abbildungen 4.2, 4.3 und 4.4 ist deutlich die schnellere Konvergenzgeschwindigkeit des Kalman-Filters gegenüber der SQP-Methode zu erkennen. Der Verlauf der Fehlergröße in Abbildung 4.5c weist keine größeren Sprünge auf und nach 27 Iterationen ist die Identifikation erfolgreich beendet. Das Ergebnis der Identifikation kann Abbildung 4.5d entnommen werden.

Im Gegensatz zur Identifikation mit Hilfe des SQP-Verfahrens hat die Wahl des Startvektors keine Auswirkungen auf das Identifikationsergebnis. In Abbildung 4.6 kann eine erfolgreiche Identifikation der Ellipse durchgeführt werden, obwohl der gleiche ungünstige Startvektor gewählt wird wie bei der SQP-Methode in Abbildung 4.4.

Identifikation von rißförmigen Fehlstellen

Die Rißerkennung verläuft genauso erfolgreich wie die Erkennung elliptischer Schädigungen. In Abbildung 4.7 sind die Ergebnisse der Identifikation von zwei Rissen zusammengestellt. Die Abbildungen 4.4 und 4.7 können direkt miteinander verglichen werden, da der gleiche Startvektor für die Optimierung verwendet wird. Das Kalman-Filter benötigt deutlich weniger Iterationsschritte zum Auffinden der gesuchten Risse als die SQP-Methode. Bei Betrachtung des Rechenaufwandes muß aber die Matrizeninvertierung bei Durchführung jedes Iterationsschrittes berücksichtigt werden. Der Rechenaufwand für einen Iterationsschritt bei Anwendung des Kalman-Filters ist auf Grund dessen größer als bei Verwendung der SQP-Methode. Bei den hier ausgewählten Beispielen war die Identifikation durch das Kalman-Filter immer schneller als die Identifikation mit dem SQP-Verfahren.

Bei Erweiterung der Identifikation auf dreidimensionale Geometrien liefert das Kalman-Filter weiterhin gute Ergebnisse. Es treten aber Probleme mit lokalen Minima auf, so daß im weiteren das Kalman-Filter nur mit einer hybriden Optimierungsstrategie im Dreidimensionalen angewandt wird.

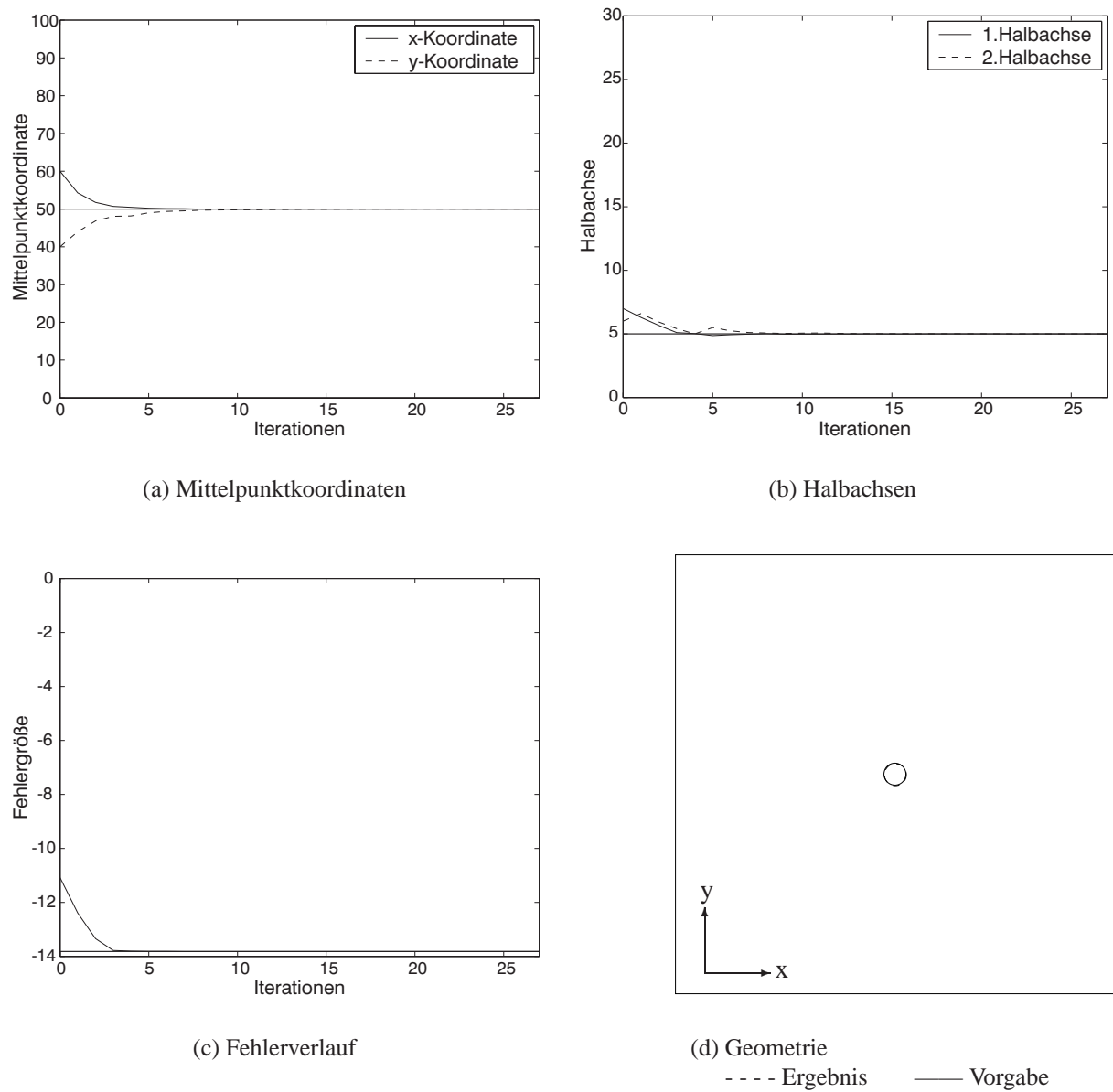
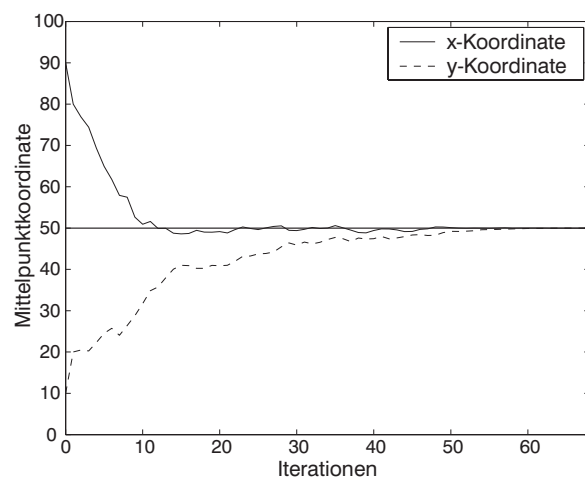
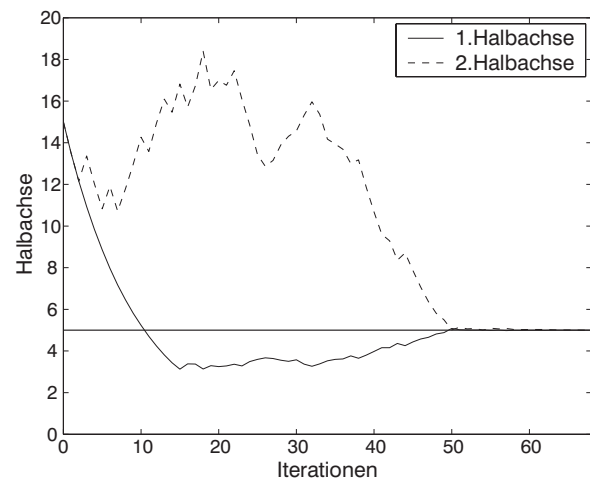


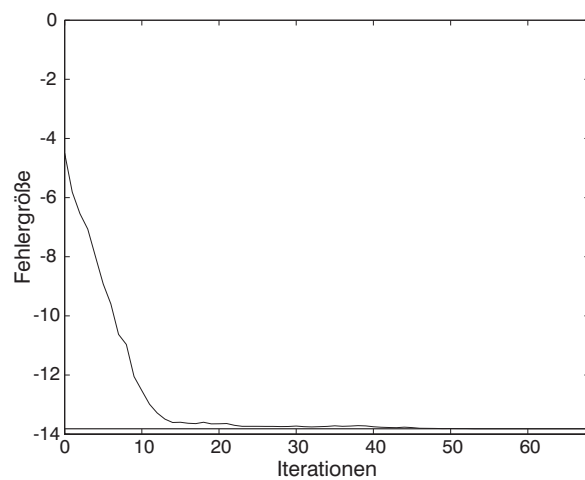
Abbildung 4.5: Identifikation einer kreisförmigen Fehlstelle bei Verwendung des Kalman-Filters mit dem Startwert [60 40 6 7]



(a) Mittelpunktkoordinaten



(b) Halbachsen



(c) Fehlerverlauf

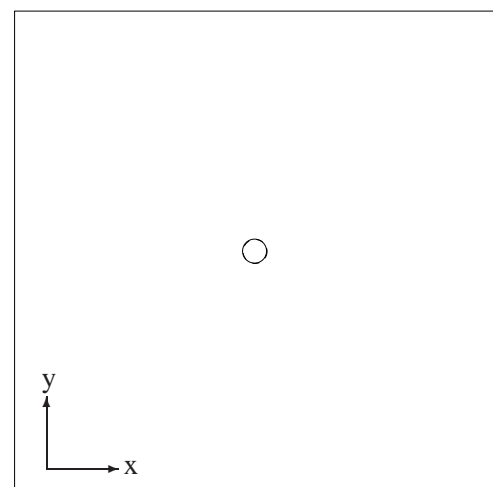
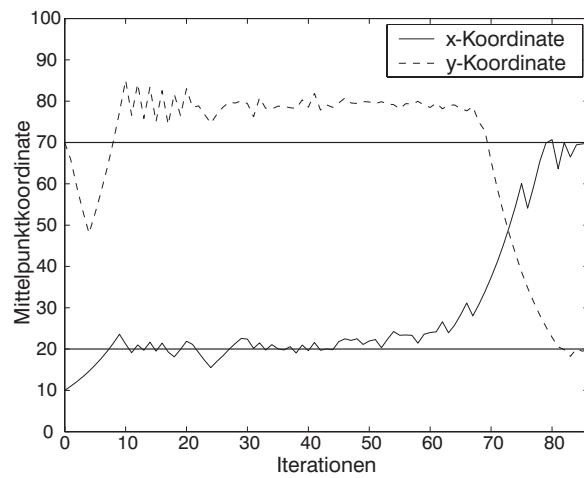
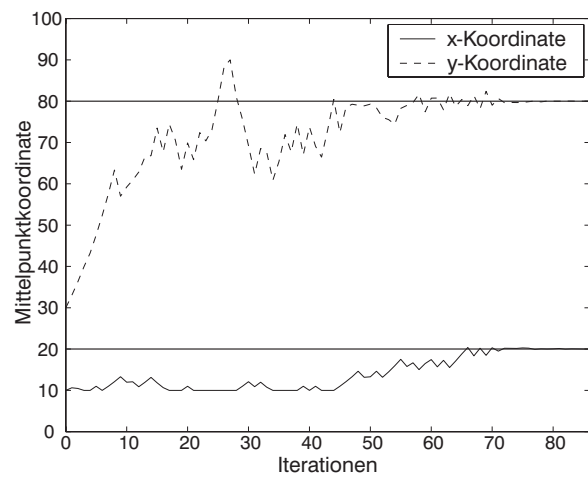
(d) Geometrie
- - - Ergebnis — Vorgabe

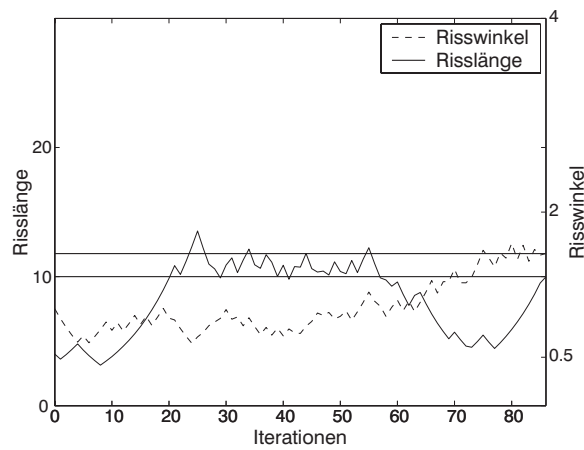
Abbildung 4.6: Identifikation einer kreisförmigen Fehlstelle bei Verwendung des Kalman-Filters mit dem Startwert [90 10 15 15]



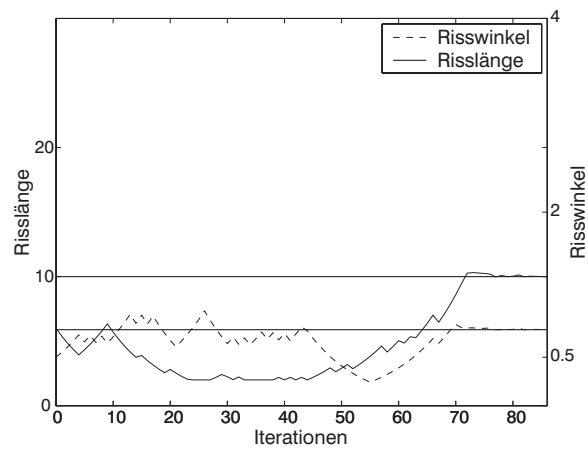
(a) Mittelpunktkoordinaten 1.Riß



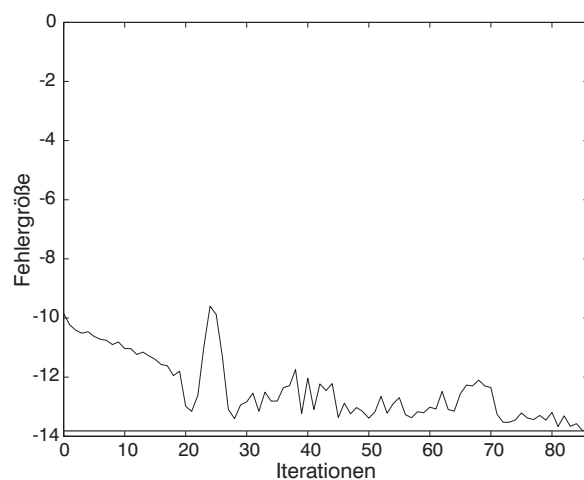
(b) Mittelpunktkoordinaten 2.Riß



(c) Rißlänge, Rißwinkel 1.Riß



(d) Rißlänge, Rißwinkel 2.Riß



(e) Fehlerverlauf

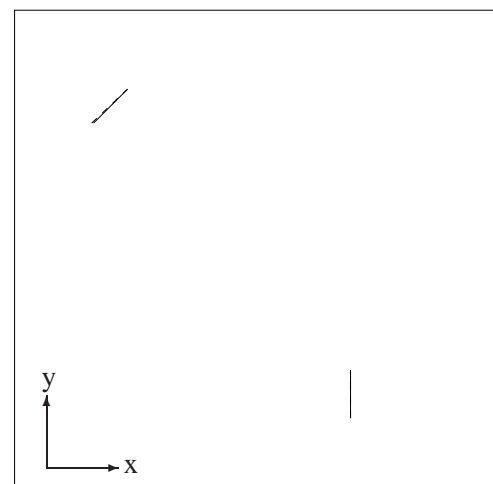

(f) Geometrie
--- Ergebnis — Vorgabe

Abbildung 4.7: Identifikation von zwei Rissen bei Verwendung des linearen Kalman-Filters

4.3 Identifikation mit Neuronalen Netzen

Die Identifikation mit neuronalen Netzen verfolgt eine andere Vorgehensweise bei der Erkennung von Fehlstellen einer Struktur als die der iterativen Optimierungsstrategien. Zunächst muß das neuronale Netz einem Lernvorgang unterworfen werden. Erst nach diesem Training kann das neuronale Netz Fehlstellen identifizieren.

Mit Hilfe der REM wird eine Scheibe mit den Abmessungen 100*100 nach Lastfall 1 in Abbildung 4.1a mit einer unter einem Neigungswinkel von ca. 6° angreifenden Streckenlast berechnet. An dieser Stelle sei vorweggenommen, daß die Identifikation mit einer zu der Oberfläche senkrecht stehenden Belastung der Scheibe nach Lastfall 1 in Abbildung 4.1a schlechtere Ergebnisse liefert als mit einer schräg angreifenden Streckenlast. Die Verschiebungen der belasteten Scheibenoberseite werden ermittelt und dem neuronalen Netz zur Verfügung gestellt. Die Diskretisierung einer Scheibenseite beschränkt sich auf fünf quadratische Elemente. Somit stehen dem Netz insgesamt 22 Verschiebungswerte als Eingangsgrößen zur Verfügung. Die Schädigung wird als kreisförmiges Loch mit einem Durchmesser von $d = 10$ vorgegeben. Wie im Abschnitt 3.4 erwähnt, wird eine Lerndatenmenge als auch eine Testdatenmenge für das neuronale Netz benötigt. Der Lernmengendatensatz besteht aus 81 Verschiebungsdatsätzen. Beginnend mit einer Schädigung bei den Koordinaten $x_c = 10$ und $y_c = 10$ werden diese in Schritten von jeweils 10 je Koordinatenrichtung erhöht. Dabei werden jeweils die Verformungen des oberen Scheibenrandes gespeichert. Der letzte Lerndatensatz beinhaltet die Verformungen der Scheibenoberseite für eine Schädigung an der Position $x_c = 90$ und $y_c = 90$. Der Testdatensatz dient zur Kontrolle der Leistungsfähigkeit des neuronalen Netzes und ist nicht im Lerndatensatz enthalten. In ihm werden die Verformungsgrößen der Scheibenoberseite für die Lochposition beginnend von $x_c = 15$ und $y_c = 15$ bis hin zu $x_c = 85$ und $y_c = 85$ in 10-er Schritten gespeichert. Bei Vorlage dieser Verformungen soll das neuronale Netz die Schadensstelle möglichst genau prognostizieren.

Die verwendeten neuronalen Netze beschränken sich auf feed-forward Netzwerkarchitekturen. Es findet somit keine Rückkopplung von Neuronen höherer Schichten zu Neuronen niedriger Schichten statt. Untersucht wird der Einfluß unterschiedlicher Lernmengen. Dabei wird sowohl die Anzahl der Lerndaten als auch die Anzahl der Eingangsneuronen variiert. Als Aktivierungsfunktion wird der tangens hyperbolicus gewählt, da dieser auch negative Werte der Neuronen berücksichtigen kann. Die Eingangsdaten werden somit auf Werte im Bereich von -0,9 bis 0,9 normiert. Desweiteren werden die Anzahl und Größe der verdeckten Schichten variiert. Damit kann entschieden auf die Leistungsfähigkeit des Netzes Einfluß genommen werden. Es werden unterschiedliche Lernstrategien untersucht. Diese wirken sich unter anderem auf die Trainingszeit und auf den Speicherbedarf der Rechnung aus.

4.3.1 Einfluß unterschiedlicher Lernstrategien

Ein wesentliches Merkmal des neuronalen Netzes ist der Lernalgorithmus mit dem das Netz trainiert wird. Der Lernalgorithmus ist entscheidend für die Qualität des trainierten Netzes und kann außerdem die notwendige Lernzeit beeinflussen. Es wird vermutet, daß auch bei der Scha-

denserkennung unterschiedliche Trainingsalgorithmen unterschiedliche Qualitäten eines neuronalen Netzes erzeugen. In diesem Abschnitt wird der Einfluß des Lernalgorithmus auf die Leistungsfähigkeit des Netzes untersucht.

Klassischer Backpropagation Algorithmus Der klassische Backpropagationlernalgorithmus, der die Werte der Wichtungen des neuronalen Netzes in Richtung des steilsten Gradienten der Fehlerfunktion ändert, wird mit unterschiedlichen Lernraten α getestet. Zusätzlich wird bei diesem Lernalgorithmus ein Momentumterm berücksichtigt, der ein Verweilen der Netziteration in einem lokalen Minimum verhindern soll. Einzelheiten zu dem klassischen Backpropagationlernalgorithmus mit Momentumterm können in [52] nachgelesen werden.

Die Anwendung des klassischen Backpropagation Lernalgorithmus auf die Problemstellung der Schadenserkennung verdeutlicht die Schwächen des Algorithmus. Die Lernfähigkeit des Netzes ist nur ausreichend und das Training dauert lange.

Beispielhaft wird das Lernverhalten eines neuronalen Netzes untersucht, das aus einer verdeckten Zwischenschicht mit acht Neuronen besteht. Abbildung 4.8 zeigt das Trainingsverhalten des neuronalen Netzes bei Veränderung der Lernrate. Der mittlere quadratische Fehler (MSE) ist über der Anzahl der Iterationsschritte aufgetragen, die das neuronale Netz für den Lernvorgang benötigt. Deutlich ist zu erkennen, daß eine kleine Lernrate die Rechenzeit und auch die Anzahl der Iterationen des Netzwerkes bei dieser Problemstellung ansteigen läßt. Nach 200.000 Iterationen wird der Lernvorgang abgebrochen. Der Trainingsvorgang dauert knapp vier Stunden. Falls die Lernrate zu groß gewählt wird, ist das Netz nicht mehr lernfähig, wie in Abbildung 4.8d zu sehen ist. Mit größerer Lernrate wird der Fehlerverlauf, bedingt durch die größere Schrittweite, hohen Schwankungen unterworfen. Tabelle 4.1 gibt zusammenfassend den Einfluß der Lernrate auf den Verlauf der Optimierung wieder.

Desweiteren ist in Tabelle 4.1 der durchschnittliche Lernfehler angegeben. Dieser gibt an, wie groß die durchschnittliche Abweichung der prognostizierten Schadensposition von der vorgegebenen entfernt liegt. Die Angaben des maximalen und des minimalen Fehlers beziehen sich auf die Prognose des trainierten Netzes durch die Testdaten. Wichtiger ist jedoch die Angabe des maximalen Fehlers. Bei Anwendung des neuronalen Netzes auf unbekannte Datensätze ist die maximale höchste Abweichung von Interesse, um eine bessere Aussage über die Lage der tatsächlichen Schädigung treffen zu können. Die maximale Abweichung bei der Bestimmung der Schadensposition ist mit knapp vier als gering einzustufen. Die Leistungsfähigkeit der trainierten Netze wird für den Lerndatensatz durch Abbildung 4.9 wiedergegeben, für den Testdatensatz durch Abbildung 4.10. In den Abbildungen wird die vorgegebene Schädigungsposition durch 'o' dargestellt und die vom neuronalen Netz prognostizierte Schadensposition durch '+'. In Abbildung 4.9d und 4.10d ist zu erkennen, daß eine Prognose mit einer Lernrate von $\alpha = 0,5$ nicht möglich ist. Die Auswertung der Testdaten zeigt bei den gut trainierbaren Netzen sehr gute Ergebnisse. Die Koordinaten der Schädigung lassen sich gut ermitteln. In Abbildung 4.10 sind die Ergebnisse dieser Prognose dargestellt.

Insgesamt kann eine gute Identifikation mit dem klassischen Backpropagation Algorithmus getroffen werden. Allerdings benötigen die Netze eine sehr lange Rechenzeit. Diese kann mit

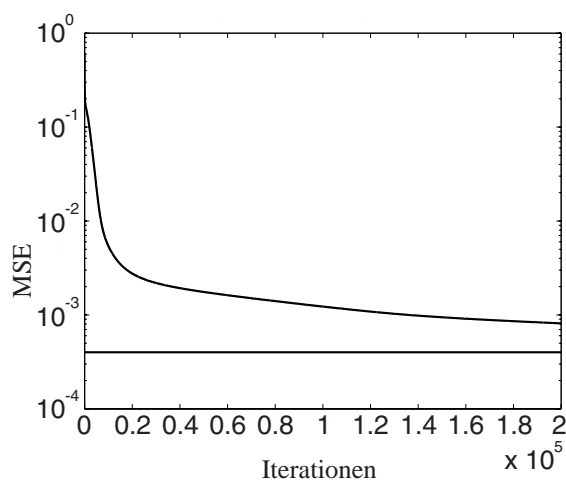
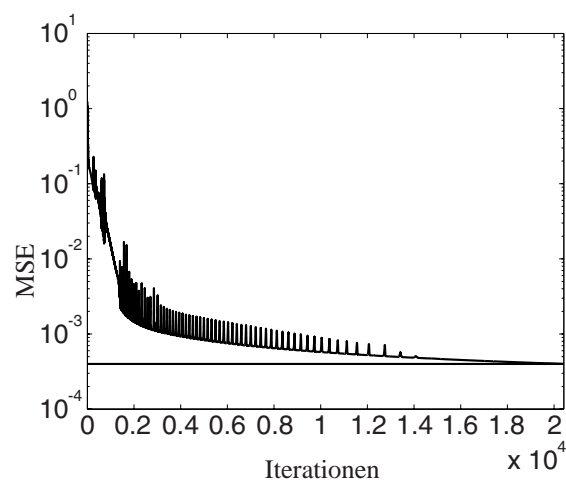
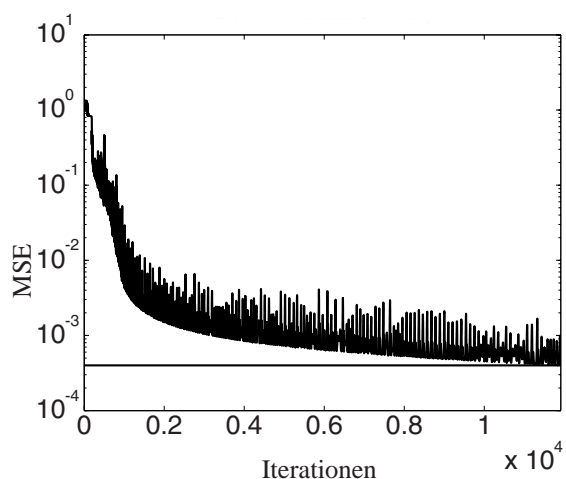
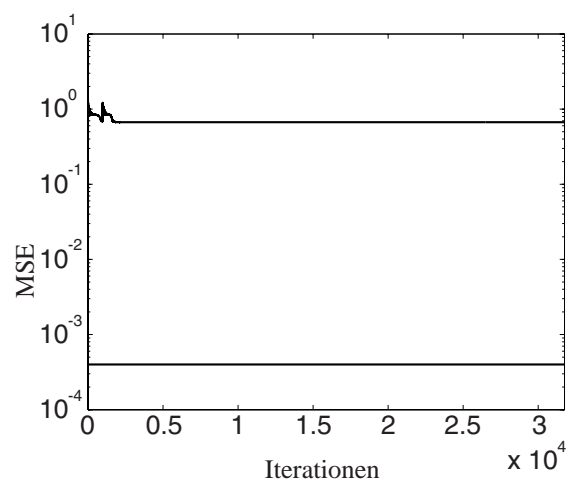
(a) Fehlerverlauf mit $\alpha = 0,1$ (b) Fehlerverlauf mit $\alpha = 0,2$ (c) Fehlerverlauf mit $\alpha = 0,4$ (d) Fehlerverlauf mit $\alpha = 0,5$

Abbildung 4.8: Fehlerverlauf während des Trainings mit unterschiedlichen Lernraten

Lernrate	Rechenzeit	Iterationen	Lernfehler ϕ	Fehler max	Fehler min
0,1	14319s	200000	1,4713	4,7814	0,1514
0,2	1034s	20421	0,9796	3,8991	0,03227
0,4	598s	11912	0,9671	3,8433	0,9671

Tabelle 4.1: Einfluß der Lernrate α auf die Rechenzeit und auf die Identifikationsergebnisse

Vorgabe einer höheren Lernrate gesenkt werden. Dabei besteht allerdings die Gefahr, daß das Netz, wie in den Abbildungen 4.9d und 4.10d zu erkennen ist, keine Ergebnisse liefert.

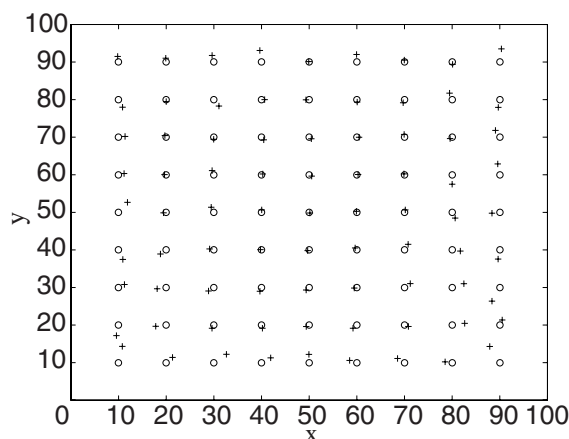
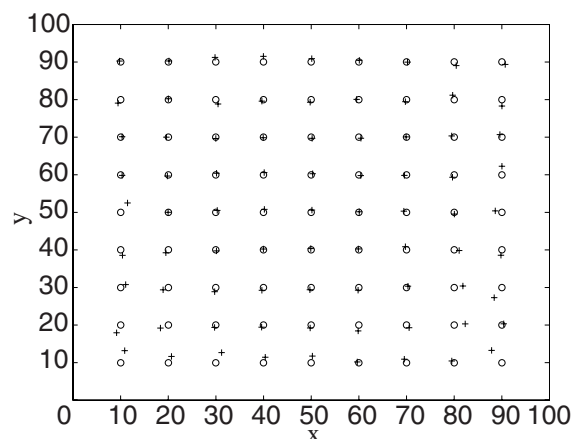
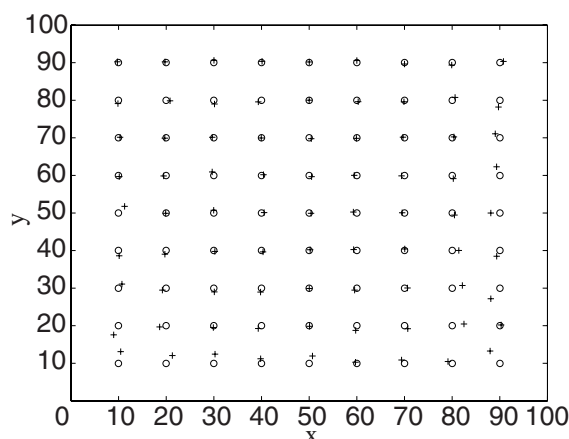
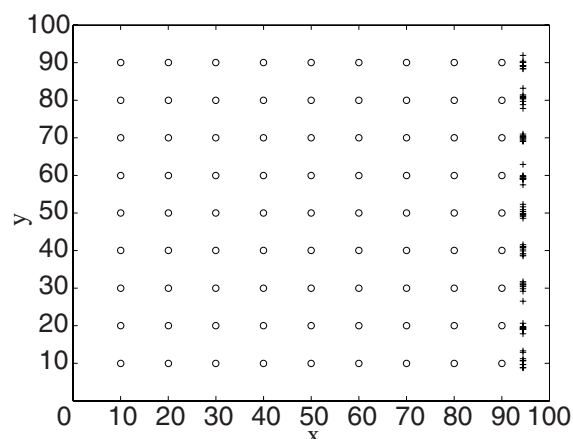
(a) Prognose der Lerndaten mit $\alpha = 0,1$ (b) Prognose der Lerndaten mit $\alpha = 0,2$ (c) Prognose der Lerndaten mit $\alpha = 0,4$ (d) Prognose der Lerndaten mit $\alpha = 0,5$

Abbildung 4.9: Prognose der Lerndaten mit unterschiedlicher Lernrate
 o vorgegebene Position + identifizierte Position

Konjugierte Gradientenverfahren Die Verwendung von konjugierten Gradientenverfahren als Lernstrategie neuronaler Netze soll die Konvergenzgeschwindigkeit der Lösung beschleunigen. Bei diesen Verfahren wird ein Gradient zur Minimierung des Fehlerwertes verwendet, der nicht dem Gradienten der steilsten Abstiegsrichtung der Fehlerfunktion entspricht. In [18] werden verschiedene Lernstrategien basierend auf konjugierte Gradientenverfahren vorgestellt. Beispielhaft wird eine Simulation mit dem „Scaled Conjugate Gradient“ Algorithmus als Lernstrategie durchgeführt. Die Eigenschaften des Algorithmus werden in [30] ausführlicher beschrieben. Abbildung 4.11 zeigt das Identifikationsergebnis. Die Prognose der Testgrößen zeigt eine gute Übereinstimmung mit den zu suchenden Parametern.

Die Verwendung weiterer Lernstrategien, die den konjugierten Gradientenverfahren zugeordnet werden, führen nicht zur erfolgreichen Identifikation. Der Lernvorgang wird vorzeitig abgebro-

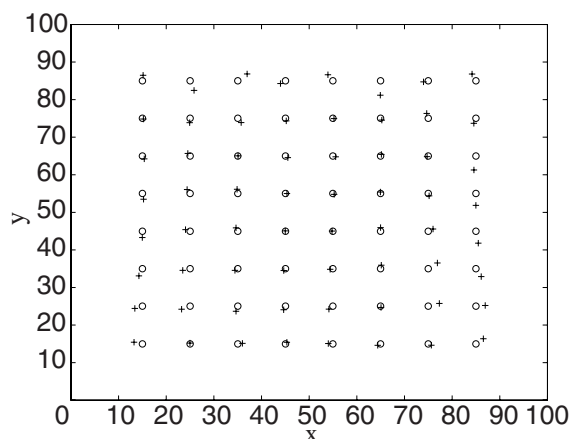
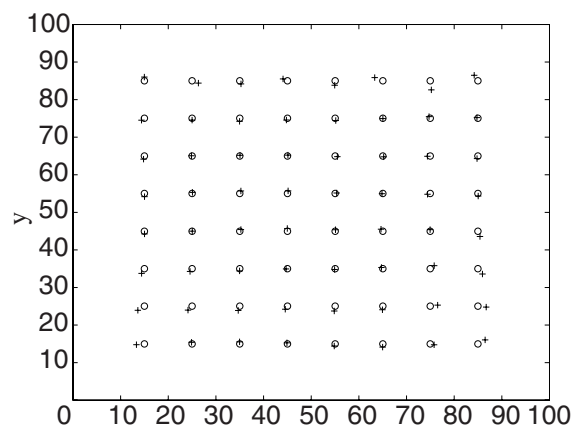
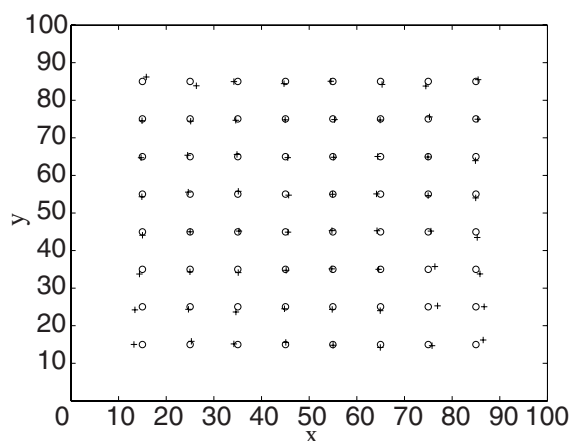
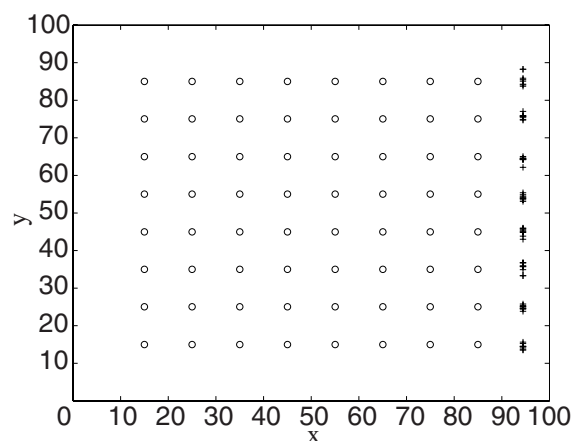
(a) Prognose der Testdaten mit $\alpha = 0,1$ (b) Prognose der Testdaten mit $\alpha = 0,2$ (c) Prognose der Testdaten mit $\alpha = 0,4$ (d) Prognose der Testdaten mit $\alpha = 0,5$

Abbildung 4.10: Prognose der Testdaten mit unterschiedlicher Lernrate
 o vorgegebene Position + identifizierte Position

chen.

Resilient Backpropagation Der Resilient Backpropagation Algorithmus (Rprop) zeichnet sich durch eine schnelle Trainingszeit aus. Anders als bei den vorher genannten Verfahren, wird bei dem Rprop Lernverfahren nicht der Wert des Gradienten selber zum Ändern der Gewichte herangezogen sondern nur dessen Vorzeichen. Weitere Eigenschaften des Algorithmus können in [34] nachgelesen werden. Ein auffälliger Nachteil dieser Lernstrategie ist, daß das Netz schnell übertrainiert wird. In diesem Falle können die Lerndaten sehr gut wiedergegeben werden, während die Testdaten sehr schlecht prognostiziert werden.

In Abbildung 4.12 ist das Ergebnis eines mit Rprop trainierten Netzes dargestellt, daß aus einer

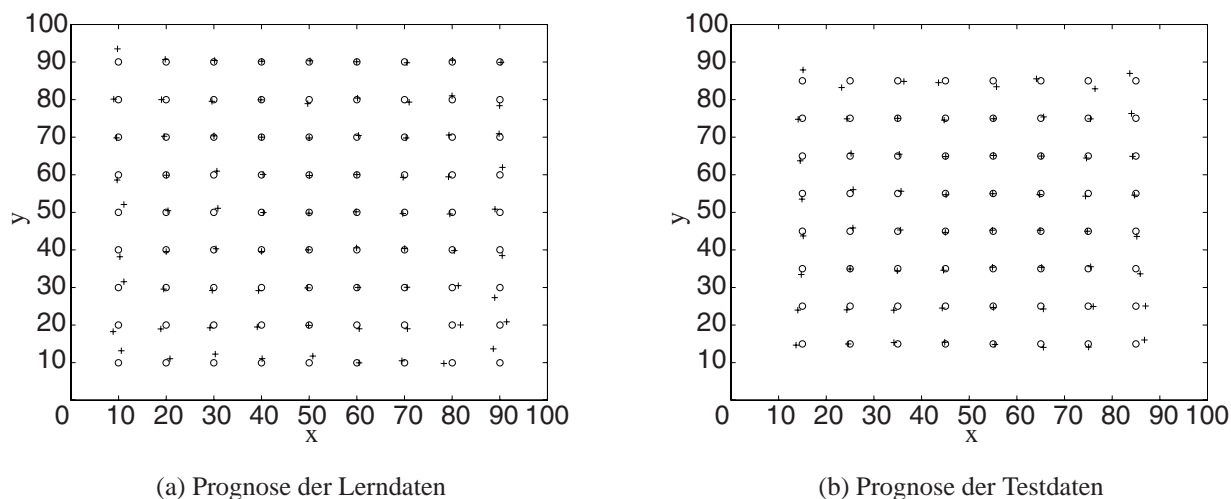


Abbildung 4.11: Identifikation mit Scaled Conjugate Algorithmus
 o vorgegebene Position + identifizierte Position

verdeckten Zwischenschicht mit 30 Neuronen aufgebaut ist. Die Prognosegenauigkeit ist im Vergleich zu anderen Testverfahren nur ausreichend. Das schnelle Übertrainieren der Netze wird auch bei diesen Untersuchungen erkannt, so daß idealerweise nach einigen Trainingsschritten der Lernvorgang kurzfristig zur Kontrolle unterbrochen werden muß. Mit den Testdaten als Eingangsgröße wird dann die Genauigkeit der prognostizierten Ergebnisse kontrolliert. Falls das Netz übertrainiert ist, werden die Prognosen im Verlauf des weiteren Trainings schlechter und das Training muß an dieser Stelle abgebrochen werden. Diese zusätzliche Untersuchung erhöht den Rechenaufwand.

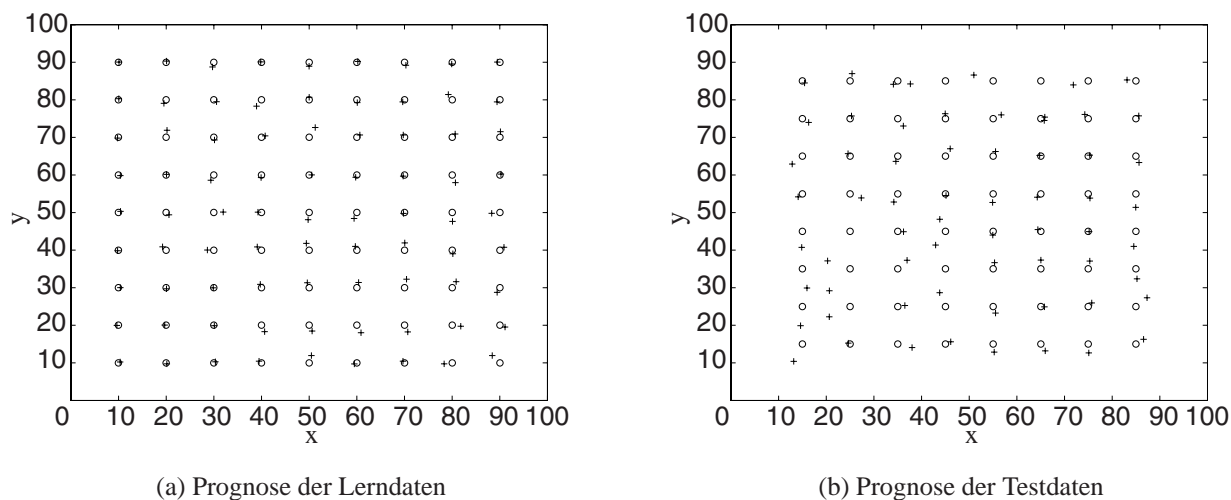


Abbildung 4.12: Identifikation mit Resilient Backpropagation Lernalgorithmus
 o vorgegebene Position + identifizierte Position

Levenberg-Marquardt Die neuronalen Netze, die mit der Levenberg-Marquardt Optimierungsstrategie trainiert worden sind, haben sich bis jetzt in der Fehlstellenidentifikation am besten bewährt. Eine akzeptable Trainingszeit und eine kleine Netzarchitektur ermöglichen gute Ergebnisse. Das Identifikationsergebnis eines mit einer verdeckten Schicht, bestehend aus acht Neuronen, angelerten Netzes kann Abbildung 4.13 entnommen werden. In [44] kann nachgelesen werden, daß selbst leicht verfälschte Meßdaten die Qualität der Identifikationsergebnisse nicht wesentlich beeinflussen. Einzelheiten zum Trainieren neuronaler Netze mit dem Levenberg-Marquardt Algorithmus können [19] entnommen werden.

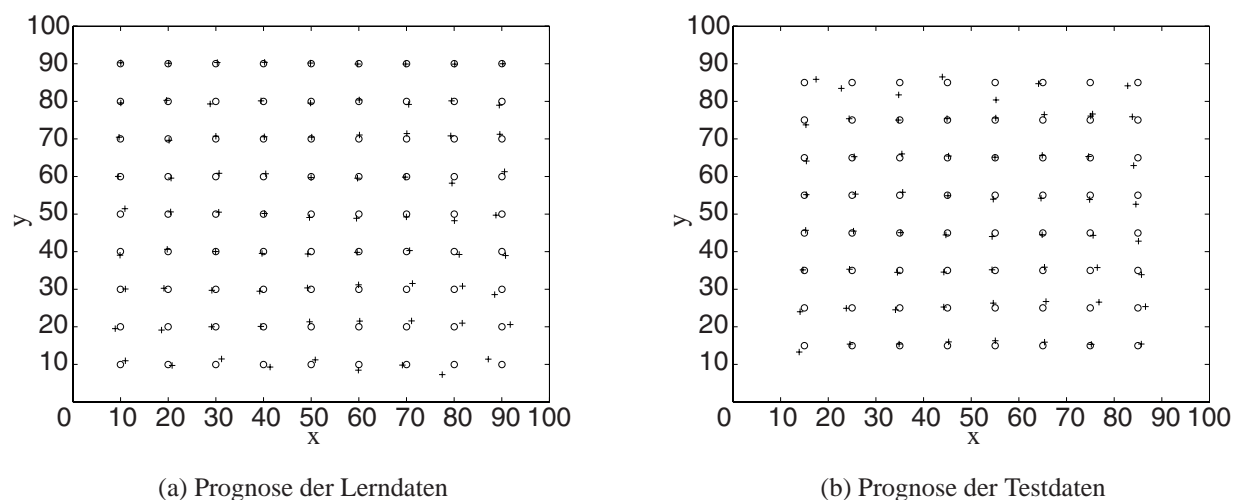


Abbildung 4.13: Identifikation mit Levenberg-Marquardt Algorithmus
o vorgegebene Position + identifizierte Position

Andere Lernstrategien Zur Schädigungslokalisierung mit neuronalen Netzen wurden weitere Trainingsstrategien getestet. Aus den Ergebnissen ergaben sich keine neuen Erkenntnisse. Unter anderem wurde das BFGS quasi Newton-Verfahren, die Bayesian Regularisierung und andere auf Newtonverfahren basierende Lernstrategien untersucht. Für eine erfolgreiche Identifikation mußte im Vergleich zu einem mit dem Levenberg-Marquardt Algorithmus trainierten Netz eine größere Netzarchitektur gewählt werden, um die gleiche Genauigkeit zu erhalten. Der Rechenaufwand stieg somit deutlich an.

4.3.2 Identifikation einer kreisförmigen Fehlstelle

Aus den gewonnenen Erkenntnissen des vorherigen Abschnittes kann die Untersuchung auf die Schadenserkennung elliptischer Fehlstellen erweitert werden. Der Radius der Fehlstelle wird dem neuronalen Netz als zusätzlich zu prognostizierende Größe hinzugefügt. Folglich muß auch die Parametrisierung der Randelements simulation angepaßt werden. Die Diskretisierung

des Randes bleibt jedoch identisch. Da nach Kapitel 4.3.1 die Levenberg-Marquardt Lernstrategie die besten Ergebnisse lieferte, wird im folgenden nur diese weiter benutzt. Das Netz besteht aus einer verdeckten Zwischenschicht mit acht Neuronen. Die numerischen Untersuchungen zeigen, daß bei Verwendung des Levenberg-Marquardt-Algorithmus kreisförmige Fehlstellen mit einem Durchmesser von $d = 4$ gut identifiziert werden können.

Identifikation mit einem Lastfall Die Identifikation bei Verwendung eines Lastfalls ergibt zufriedenstellende Ergebnisse. In Abbildung 4.14 ist zum einen die Prognose der gelernten Daten und die Prognose der Testdaten dargestellt. In Abbildung 4.14b ist deutlich zu erkennen, daß eine Identifikation der nahe an der belasteten Scheibenoberseite liegenden Fehlstellen nicht möglich ist. Um diese Fehleinschätzung des trainierten Netzes zu beseitigen, wird die Anzahl der Lerndaten zunächst erhöht. Diese Maßnahme hat unwesentlichen Einfluß auf die Genauigkeit der Ergebnisse im besagten Bereich. Die beste Lösung ist, die Prognose der Fehlstellen mit zwei Lastfällen durchzuführen.

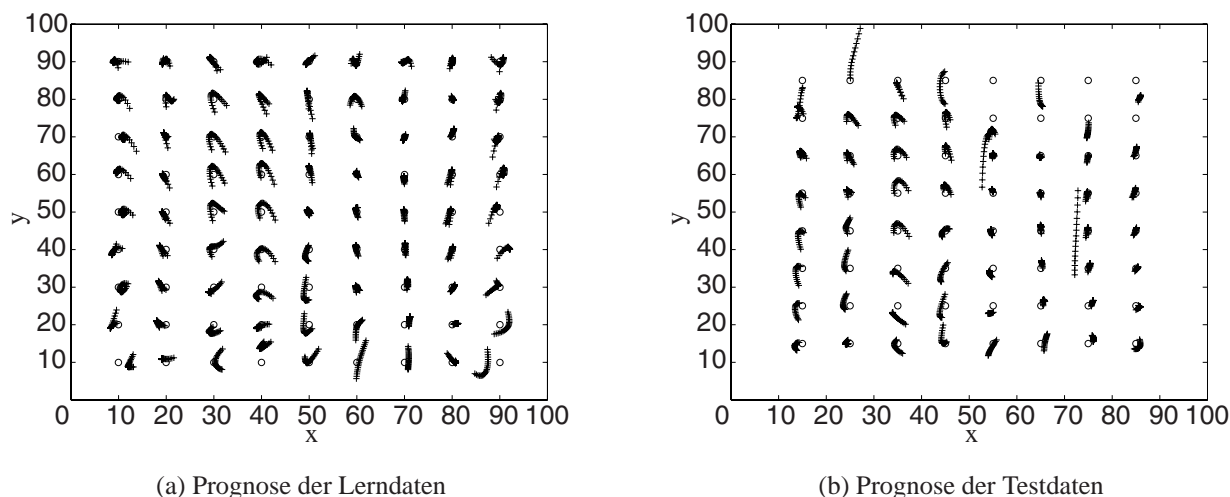


Abbildung 4.14: Identifikation der Mittelpunktkoordinaten lochförmiger Schädigungen mit einem Lastfall

o vorgegebene Position + identifizierte Position

Identifikation mit zwei Lastfällen Die Verwendung eines zusätzlichen Lastfalls bei der Identifikation der Fehlstellen führt, wie Abbildung 4.15 zeigt, auf eine Verbesserung der Genauigkeit der Prognose nahe der Scheibenoberseite. Bei diesen numerischen Untersuchungen wurden allerdings die Anzahl der Eingangsneuronen verdoppelt, da jetzt zu den 22 Verschiebungswerte aus einem Lastfall noch 22 weitere Verschiebungswerte aus dem zweiten Lastfall als Eingangsgrößen für das neuronale Netz berücksichtigt werden müssen. Es müssen somit mehr unbekannte Gewichte berechnet werden, wodurch der Speicherbedarf des Netzes und auch die benötigte Rechenzeit ansteigt. Ein weiterer Vorteil einer Einbeziehung des zweiten Lastfalls zur Prognose

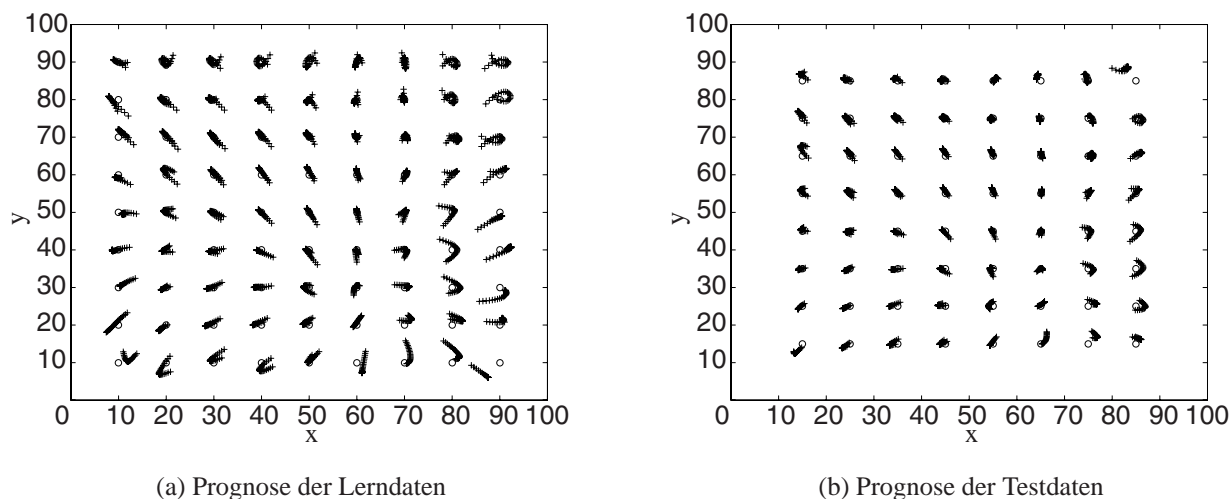


Abbildung 4.15: Identifikation der Mittelpunktkoordinaten lochförmiger Schädigungen
 o vorgegebene Position + identifizierte Position

se ist die bessere Identifikation des Halbmessers der Fehlstelle. Dieser kann nach Abbildung 4.16 mit einer geringeren Streuung angegeben werden, als wenn die Prognose nur mit einem Lastfall durchgeführt wird. In einer für die Ergebnisdarstellung neuronaler Netze üblichen Abbildung 4.16 ist die Vorgabe des Radius der Prognose des neuronalen Netzes gegenübergestellt. Stimmen die Prognosen des neuronalen Netzes mit den Vorgaben überein, befinden sich die zugehörigen Radien auf der Winkelhalbierenden der Grafik wieder. Aus den durchschnittlichen Abweichungen der Prognose zu der Vorgabe läßt sich eine Ausgleichsgerade bestimmen. Die Lage der Ausgleichsgerade ist somit ein Maß für die Qualität des trainierten neuronalen Netzes. An Hand der Streuung zeigt Abbildung 4.16 aber auch, daß die Erkennung des Radius nicht so erfolgversprechend ist wie die Lokalisation des Schadensmittelpunktes.

4.3.3 Identifikation von ellipsenförmigen Fehlstellen

Das neuronale Netz wird in seiner Architektur so geändert, daß elliptische Fehlstellen prognostiziert werden können. Durch die Hinzunahme des vierten Neurons in der Ausgabeschicht steigt die Anzahl der zu bestimmenden Wichtungen innerhalb des Netzes weiter an. Die Prognose der nun vier Unbekannten einer Schädigung gelingt nur mit größeren Abweichungen von deren Sollwerten, wenn nur zwei Lastfälle berücksichtigt werden. Ein dritter Lastfall wird hinzugenommen. Dieser Lastfall ist dem in Abbildung 4.1b sehr ähnlich. Die betragsmäßig gleiche Belastung greift an der linken Scheibenseite an, während die rechte Scheibenseite eingespannt ist. Als positiv erweist sich, die Anzahl der Lernmenge zum Trainieren des Netzes zu erhöhen. Es stehen nun 225 Mittelpunktkoordinaten von Ellipsen zur Verfügung.

In Abbildung 4.17 ist das Ergebnis der Identifikation dargestellt. Das Erkennen der Schädigungsmittelpunkte ist gut. Sowohl die Lerndaten in Abbildung 4.17a als auch die Testdaten

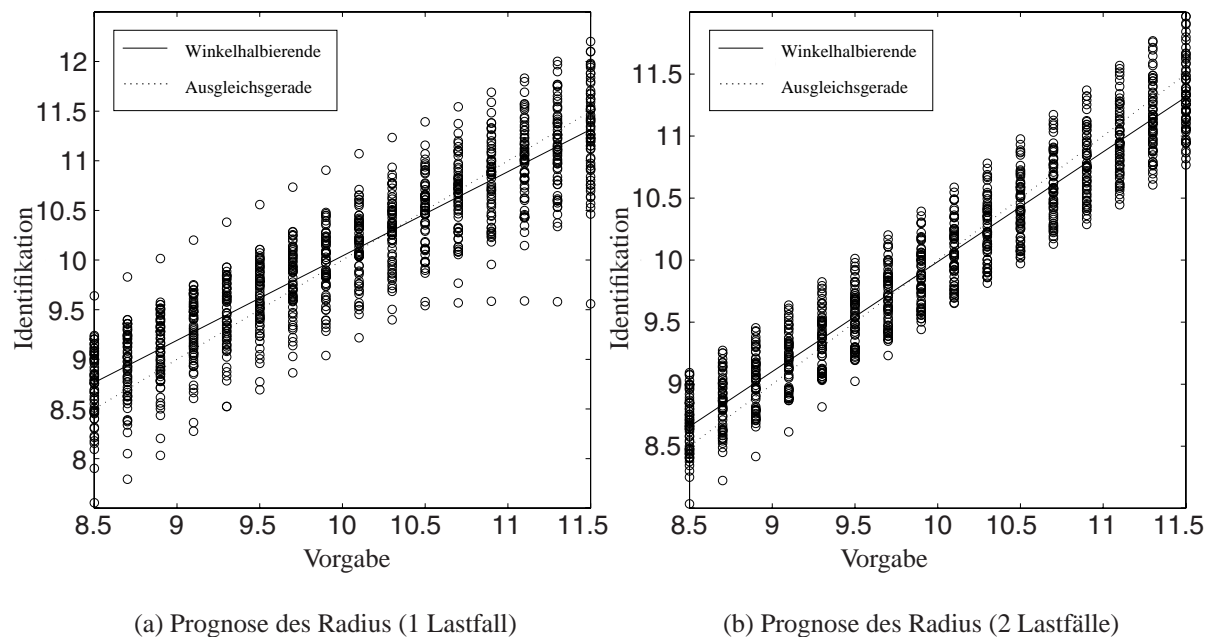


Abbildung 4.16: Identifikation des Radius lochförmiger Schädigungen mit zwei Lastfällen

in Abbildung 4.17b werden gut wiedergegeben. Eine andere Darstellung der Identifikationsergebnisse ist für die Lern Daten in Abbildung 4.18 und für die Testdaten in Abbildung 4.19 zu sehen. Die Prognose der einzelnen zu identifizierenden Größen ist den zu suchenden Parametern gegenübergestellt. Bei einer zufriedenstellenden Parameteridentifikation befinden sich die berechnenden Punkte auf der Winkelhalbierenden wieder. Der gefundene Wert entspricht dann genau der zu identifizierenden Größe. Die Identifikation der Mittelpunktkoordinaten ist in den Abbildungen 4.18a und 4.18b für die Lernmenge dargestellt. Die berechneten Punkte liegen sehr nahe an der Winkelhalbierenden. Das gleiche gilt für die Identifizierung der Koordinaten mit dem Testdatensatz in Abbildung 4.19a und 4.19b. Die Abbildungen 4.18c und 4.18d verdeutlichen aber die Ungenauigkeiten bei der Identifikation der Halbachsen der elliptischen Schädigungen schon nach dem Trainingsvorgang des Netzes. Die Streuung der prognostizierten Werte in Abhängigkeit der vorgegeben Größen ist zum Teil größer als 15%. Die Identifikation der Radien mit dem Testdatensatz, abgebildet in den Abbildungen 4.19c und 4.19d liefert ebenfalls ungenaue Ergebnisse. Es kann nur eine ungefähre Größenordnung der Schädigung angegeben werden. Auch die Erweiterung des neuronalen Netzes auf zwei verdeckte Zwischenschichten sowie eine Vergrößerung der Neuronenanzahl in diesen Schichten hat nur einen unwesentlichen Einfluß auf die Verbesserung der Identifikationsgenauigkeit. Eine Ausweitung der Studien auf mehrerer Schädigungen innerhalb einer Scheibe wird nicht weiter verfolgt, da der Aufwand für die Erstellung der Lern- und Testdatensätze stark anwächst. Zudem ist der zeitliche Aufwand, der für den Lernvorgang eines neuronalen Netzes benötigt wird, sehr hoch.

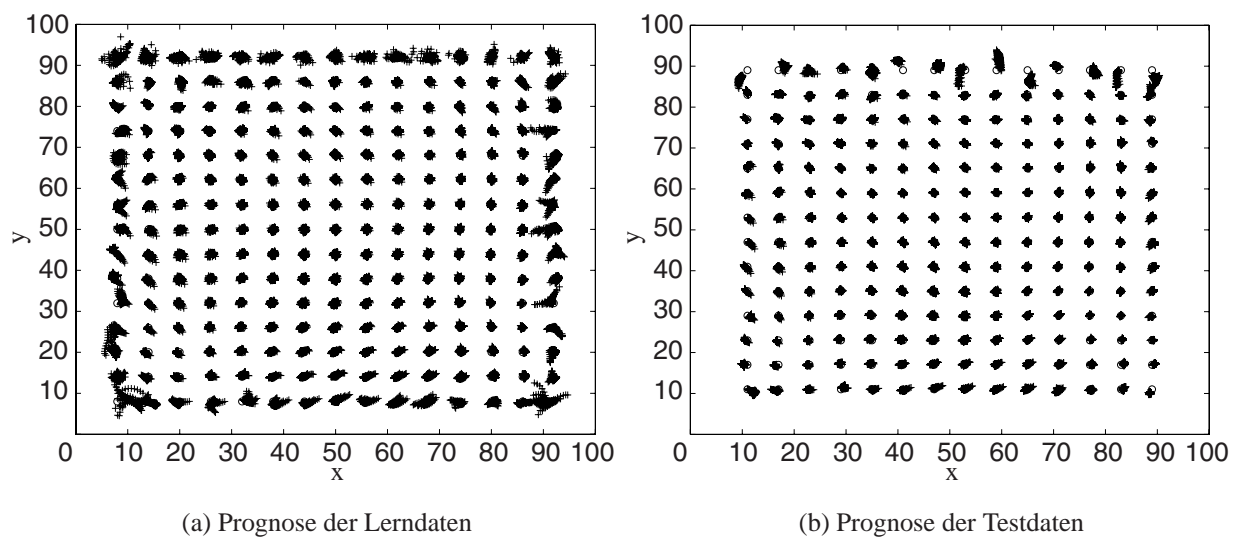
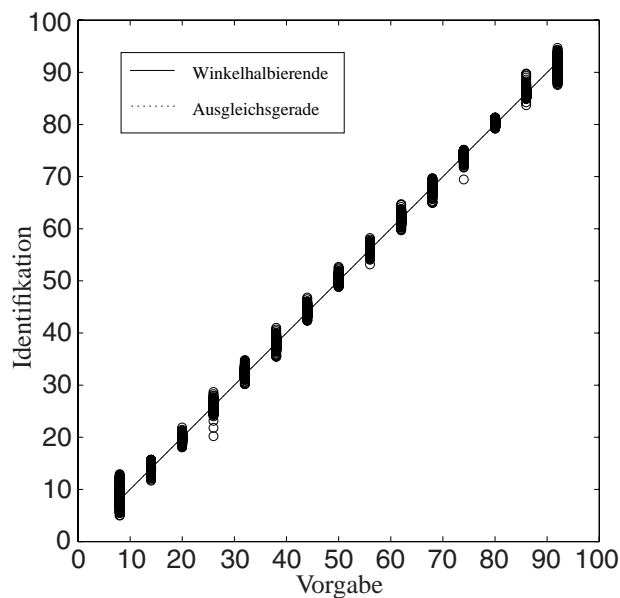
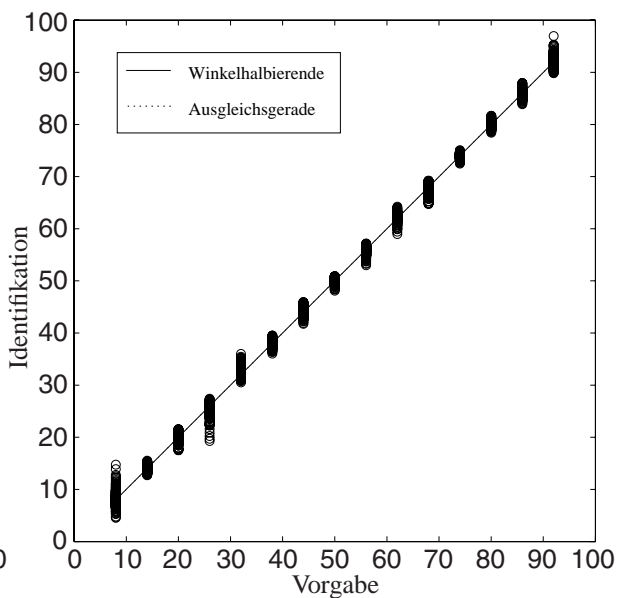


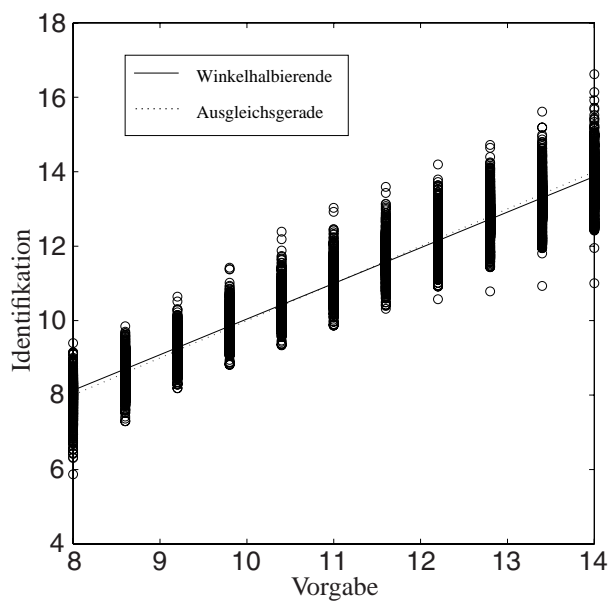
Abbildung 4.17: Identifikation der Mittelpunktkoordinaten mit Levenberg-Marquardt Strategie und 225 Lerndatensätze
 o vorgegebene Position + identifizierte Position



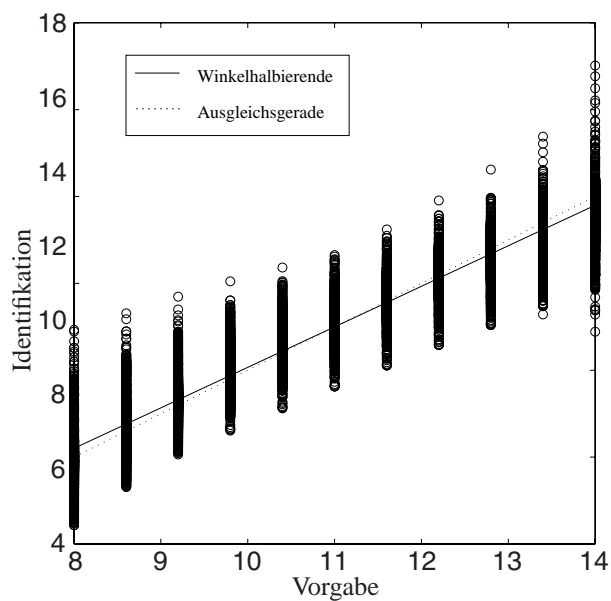
(a) Prognose der x-Koordinate



(b) Prognose der y-Koordinate



(c) Prognose der 1. Halbachse



(d) Prognose der 2. Halbachse

Abbildung 4.18: Identifikation der Parameter elliptischer Fehlstellen mit den Lerndaten

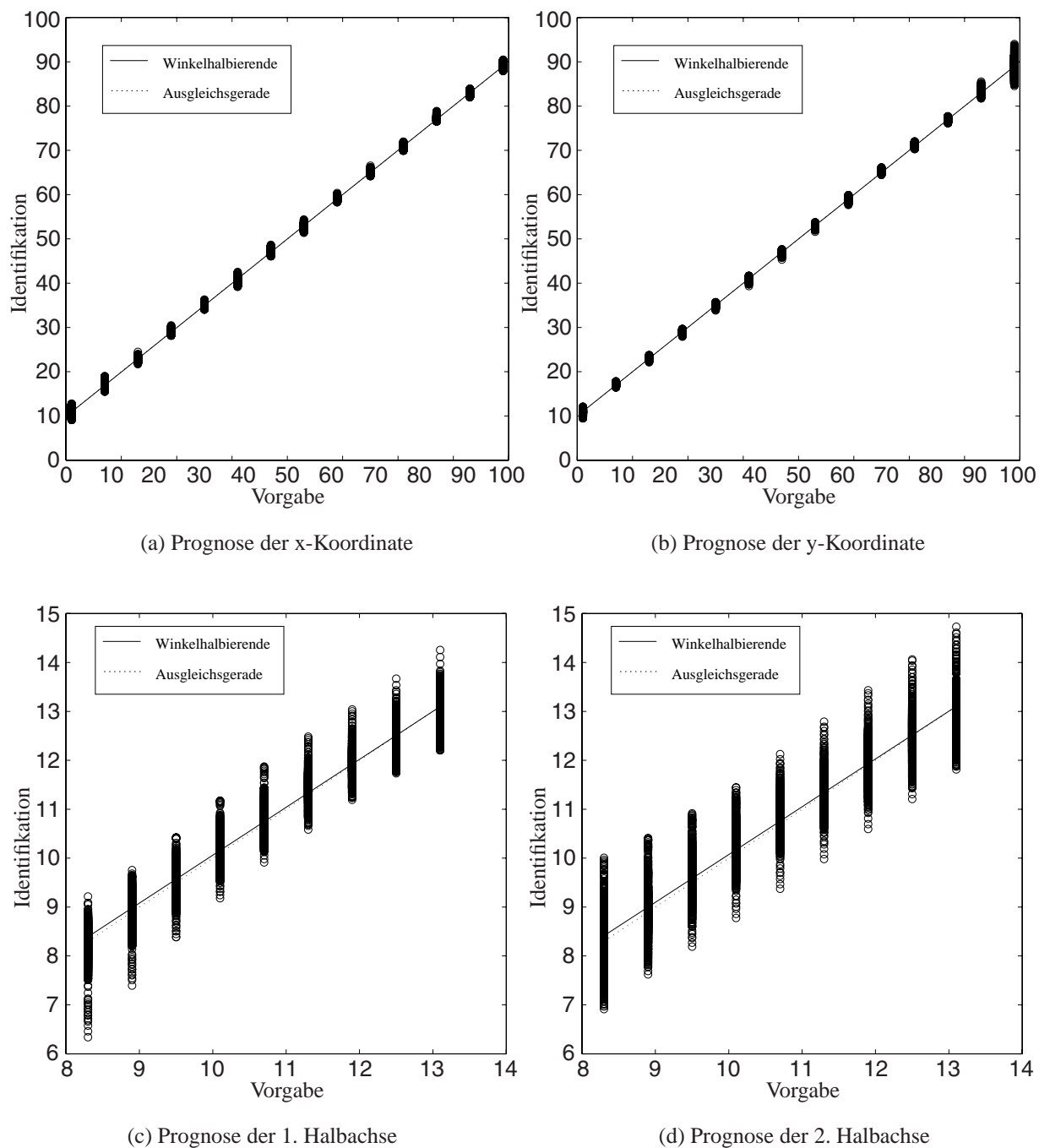


Abbildung 4.19: Identifikation der Parameter elliptischer Fehlstellen mit den Testdaten

4.4 Identifikation mit dem genetischen Algorithmus

Der genetische Algorithmus wird so implementiert, daß er eine quadratische Fehlerfunktion

$$E(\mathbf{z}) = \sum_{j=1}^{2n} (u_j(\mathbf{z}) - u_{0j}(\mathbf{z}_0))^2 \quad (4.2)$$

minimiert. Diese Fehlerfunktion, auch Fitneßfunktion genannt, stellt das Fehlerquadrat der Verformungen der Scheibenaußenseite dar. In Gleichung (4.2) stellt u_0 die Verformungsgrößen des Experimentes dar, während u die Verformungsgrößen zum derzeitigen Iterationsschritt beinhaltet. n gibt die Anzahl der ausgewählten Meßpunkte an. Je Meßpunkt werden die Verschiebungen der Scheibenaußenseite in beide Koordinatenrichtungen für die Berechnung der Fehlerfunktion herangezogen.

Ein Abbruchkriterium muß vorgegeben werden, damit bei Erreichen einer bestimmten Genauigkeit der Ergebnisse die Rechnung unterbrochen wird. Folgende Abbruchkriterien werden bei diesem genetischen Algorithmus verwendet:

- Bei einer maximalen Anzahl von 100.000 durchgeführten Iterationsschritten wird die Optimierung abgebrochen.
- Der Abbruch erfolgt nach 10.000 erfolglosen Iterationsschritten, wenn keine Verbesserung des Fitneßwertes erzielt worden ist.
- Der Abbruch der Optimierung erfolgt bei Unterschreitung eines von der Aufgabenstellung abhängigen Wert des Fitneßwertes.

Die Festlegung eines Fitneßwertes als untere Schranke der Optimierung erfolgt problemspezifisch. Es müssen zunächst einige Testrechnungen durchgeführt werden, um eine Aussage über die Qualität der Ergebnisse in Abhängigkeit des Fitneßwertes treffen zu können.

Bei diesen Testbeispielen wurden fünf lineare Elemente je Scheibenseite verwendet und die elliptischen Fehlstellen mit zehn linearen Elementen diskretisiert. Um Rechenzeit einzusparen, wurde auf eine feinere Diskretisierung verzichtet. Bevor die Fehlstellenidentifikation gestartet werden kann, müssen die einzelnen Parameter des genetischen Algorithmus festgelegt werden. Für die Berechnungen wird eine Populationsgröße von 100 Individuen gewählt. Nach jeder Iteration wird den einzelnen Chromosomen ihr jeweiliger Fitneßwert zugewiesen. Diese werden durch ein Selektionsverfahren [38] nach ihrem Zielwert beurteilt. Die 70 Individuen mit den besten Fitneßwerten bleiben der Generation erhalten, damit müssen nach jeder Iteration 30 neue Chromosomen erstellt werden. Die einzelnen Chromosomen bestehen aus vier Parametern, die jeweils die Schadensposition und die Halbachsen der elliptischen Schädigung beinhalten. Die erste Generation wird zufällig erzeugt, wobei die Grenzen der zu optimierenden Größen berücksichtigt werden müssen.

Damit zuverlässige Ergebnisse mit der Randelementmethode berechnet werden können, muß darauf geachtet werden, daß der genetische Algorithmus Werte für die Schädigung innerhalb

eines zulässigen Bereiches wählt. Der Mittelpunkt der elliptischen Schädigung muß innerhalb der Scheibe liegen. Daher wird die Nebenbedingung so formuliert, daß der Mittelpunkt nur in einem Intervall zwischen fünf und 95 im globalen Koordinatensystem der Scheibe variiert werden kann. Das globale Koordinatensystem hat in der unteren linken Ecke der Scheibe seinen Ursprung. Die Halbachsen haben eine minimale Größe von eins und eine maximale von 30. Zur Erzeugung neuer Individuen kommt das Zwei-Punkt-Crossover [38] zur Anwendung. Die Mutation der einzelnen Gene erfolgt nach der Gauß-Verteilung.

4.4.1 Identifikation von elliptischen Schädigungen mit dem genetischen Algorithmus in einer Scheibe

In diesem Abschnitt wird zuerst der Einfluß der Lastanzahl auf den Verlauf der Optimierung untersucht. Es wird erwartet, daß bei Verwendung mehrerer Lastfälle die Schadenserkennung schneller durchgeführt werden kann. Darauf aufbauend, wird untersucht, in wie weit der genetische Algorithmus Schädigungen identifizieren kann, wenn eine falsche Anzahl an Schädigungen vorgegeben wird. Die Optimierung sollte in diesem Fall nicht erfolgreich beendet werden können. Abschließend wird der Optimierung die Anzahl der Schädigungen als eine weitere Unbekannte hinzugefügt. Damit muß der genetische Algorithmus neben den Schädigungsparametern auch die Anzahl der Schädigungen selbst identifizieren.

Einfluß der Lastfallanzahl Ein Riß oder zwei Risse konnten mit einem Lastfall immer gefunden werden [12]. Deswegen wird an dieser Stelle nicht weiter darauf eingegangen. Der Einfluß mehrerer Lastfälle soll untersucht werden. Es kann vermutet werden, daß der Optimierungsverlauf und damit die Identifikation der Schädigungsparameter auf diese Weise verbessert werden kann. Um dies zu überprüfen, wird das inverse Problem auf drei zu identifizierende Ellipsen erweitert. Die Lage der Fehlstellen wird so gewählt, daß eine Ellipse direkt unter einer anderen liegt. Bei senkrecht zur Scheibenoberseite wirkender Belastung überlagern sich die Verformungsauswirkungen der beiden Fehlstellen und verhindern so eventuell die Identifikation. Die Anordnung der Fehlstellen kann Abbildung 4.20 entnommen werden. Da die Verformungen beider Lastfälle die gleiche Größenordnung haben, werden zur Optimierung die Fehlerfunktionen beider Lastfälle addiert. Bei der Untersuchung mit einem Lastfall wird diese Fehlerfunktion verdoppelt, um die Fitneßfunktionen beider Optimierungsvorgänge vergleichen zu können.

Ein Lastfall Der genetische Algorithmus liefert mit nur einem Lastfall schlechte oder in diesem Fall falsche Ergebnisse. Die Suche wird nach 80.000 Iterationen abgebrochen, da sich bei den letzten 10.000 Iterationen der Fitneßwert nicht wesentlich verändert hat. Der Fitneßwert der letzten Iteration beträgt 6,0517 und ist damit um den Faktor 60 höher als der geforderte Grenzwert. Die Rechenzeit für diese Simulation beträgt nahezu vier Tage.

Die Auswertung der Ergebnisse in Abbildung 4.20a zeigt eine gute Annäherung einer Ellipse, die gefundenen Parameter der anderen Schädigungen liefern jedoch unbrauchbare Ergebnisse. Wie erwartet, beeinflussen die in Lastrichtung übereinander positionierten Ellipsen den Opti-

mierungsverlauf negativ. Die Verwendung von nur einem Lastfall ist nicht ausreichend für eine Identifikation zweier Schädigungen, deren Verformungsauswirkungen sich überlagern.

Zwei Lastfälle Eine erhebliche Verbesserung der Genauigkeit und der Verkürzung der Identifikationszeit ergibt sich bei der Berücksichtigung eines zweiten Lastfalls. Schon nach 18.100 Iterationsschritten ist der Grenzwert der Fitneßfunktion mit 0,099 unterschritten. Die vom Algorithmus benötigte Rechenzeit beträgt knapp zwei Tage.

Die elliptischen Schädigungen werden, wie Abbildung 4.20b zeigt, gut identifiziert. In Abbildung 4.21 wird die Verbesserung des Optimierungsverlaufes auf Grund der Verwendung eines zweiten Lastfalls deutlich. Der Fitneßwert fällt während der Optimierung sehr deutlich ab. Die Identifikation des dritten Loches wird trotz dessen ungünstiger Lage durch den zweiten Lastfall ermöglicht.

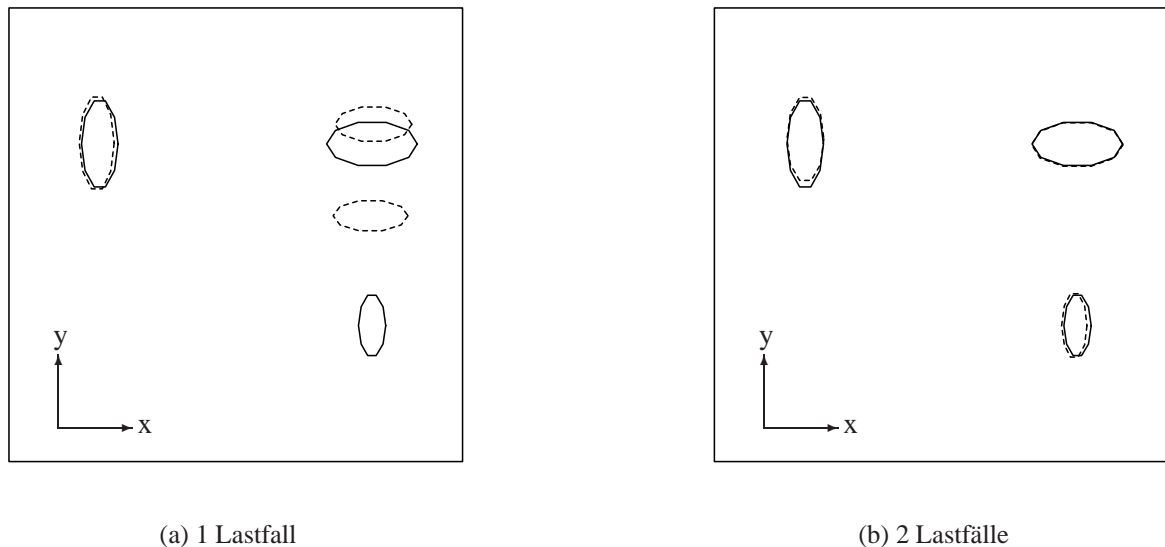


Abbildung 4.20: Identifikationsergebnis von drei elliptischen Schädigungen:

Vergleich Ein Lastfall → Zwei Lastfälle

- - - Ein Lastfall ——— Zwei Lastfälle

Identifikation mit richtig vorgegebener Lochanzahl Die Ergebnisse aus dem vorherigen Abschnitt zeigen, daß es sinnvoll ist, weitere Berechnungen mit zwei Lastfällen durchzuführen. In dieser Simulation werden fünf Ellipsen in die Scheibe integriert. Da der genetische Algorithmus bisher immer in der Lage war, gleich große Schädigungen zu lokalisieren, wird in dieser Rechnung bewußt eine kleinere Schädigung berücksichtigt. Die Anordnung der Schädigungen innerhalb der Scheibenstruktur kann Abbildung 4.22b entnommen werden. Nach 10.000 Iterationen wird die Simulation nach einer Rechenzeit von drei Tagen, vier Stunden und 40 Minuten beendet, da davon ausgegangen werden kann, daß der Fitneßwert den unteren Grenzwert nicht

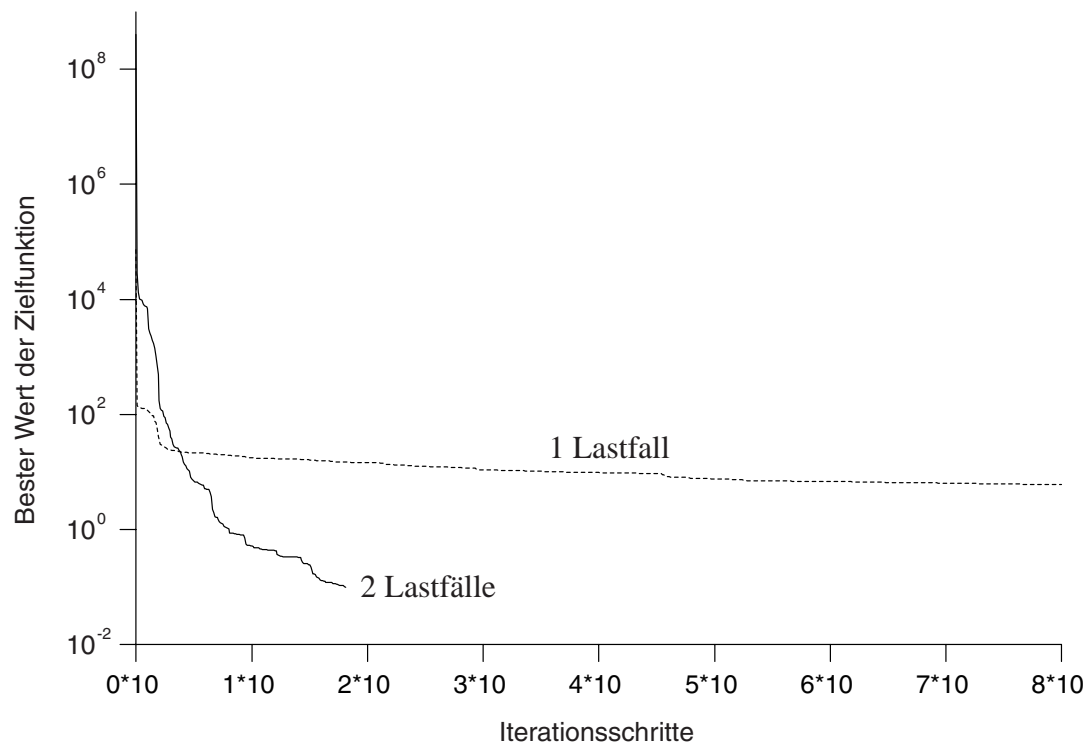


Abbildung 4.21: Fitneßverlauf bei der Schadenserkennung von drei Ellipsen: Vergleich ein Lastfall → zwei Lastfälle

erreichen kann. Der Verlauf der Fitneßfunktion ist in Abbildung 4.22a dargestellt. Das Ergebnis der Simulation zeigt, daß der genetische Algorithmus nicht in der Lage ist, eine kleine Schädigung bei Vorhandensein einer größeren zu identifizieren. Der schlechte Fitneßwert deutet diesen Sachverhalt schon während der Optimierung an. Die anderen vier Schädigungen können mit einigen Ungenauigkeiten gut erkannt werden.

Identifikation mit falsch vorgegebener Lochanzahl Das Ziel der folgenden Untersuchung ist es, die Auswirkung einer dem genetischen Algorithmus falsch vorgegebenen Anzahl an Schädigungen auf die Identifikationsergebnisse festzustellen. Meistens ist nicht nur die Lage und die Form der Schädigung unbekannt, sondern auch die Anzahl der vorhandenen Schädigungen.

Als Beispiel wird eine inverse Problemstellung berechnet, in der vier Fehlstellen in einer Scheibe vom genetischen Algorithmus identifiziert werden sollen. Allerdings wird dem Algorithmus vorgegeben, daß er nur zwei Schädigungen identifizieren soll. Die geometrische Lage und Größe der Defekte kann Abbildung 4.23b entnommen werden. Nach 30.000 Iterationsschritten wird der Algorithmus abgebrochen, da sich nach Abbildung 4.23a der Fitneßwert über die letzten 20.000 Iterationsschritte nicht verbessert hat. Der sehr schlechte Fitneßwert von 17896,9 läßt aber schnell erkennen, das eine falsche Anzahl an Schädigungen in den Voreinstellungen des

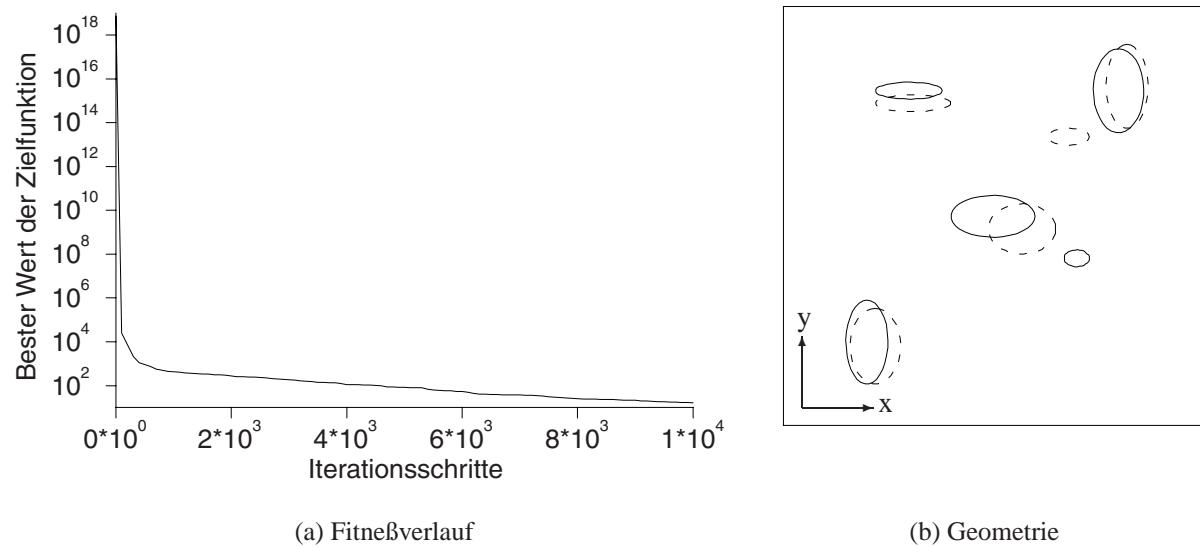


Abbildung 4.22: Identifikationsergebnis von fünf elliptischen Schädigungen
 - - - Ergebnis ——— Vorgabe

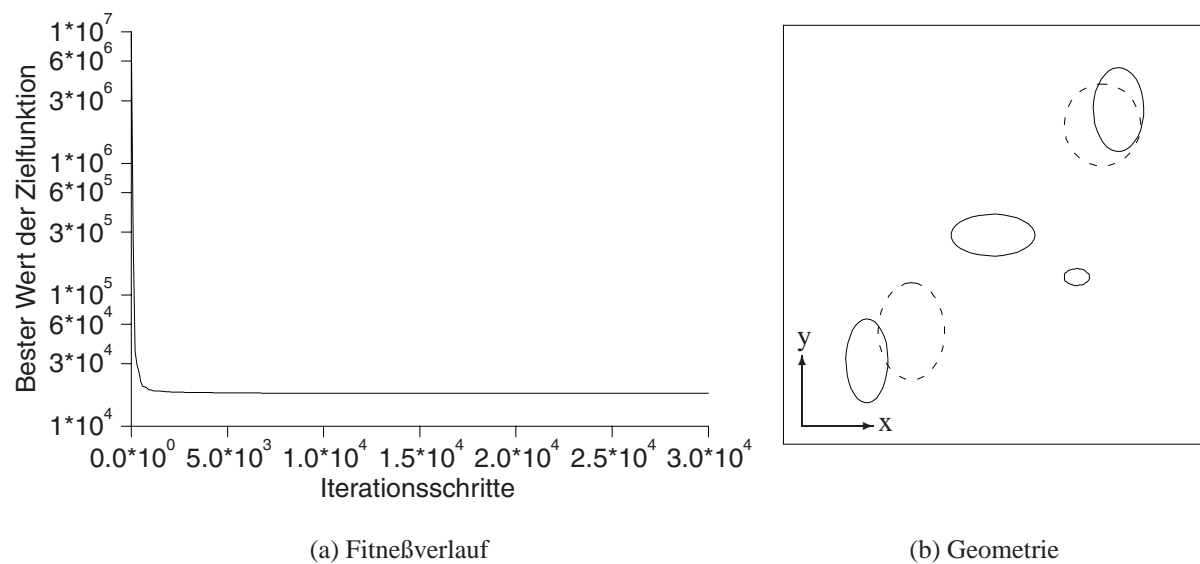


Abbildung 4.23: Identifikationsergebnis von vier elliptischen Schädigungen bei falsch vorgegebener Anzahl von zwei Defekten
 - - - Ergebnis ——— Vorgabe

Algorithmus gewählt wurde.

Identifikation bei unbekannter Lochanzahl Da im allgemeinen die Anzahl der vorhandenen Schädigungen nicht bekannt ist, wird diese als eine weitere Unbekannte dem genetischen Algorithmus zur Verfügung gestellt. Um diesen Parameter richtig zu berücksichtigen, muß innerhalb des genetischen Algorithmus eine Ganzzahloptimierung eingeführt werden. Es kann eine beliebige Anzahl von einer bis zu drei Ellipsen in der Scheibe ermittelt werden, wobei die Obergrenze der maximal auftretenden Schädigungen vor Beginn der Optimierung festgelegt werden muß. Die Lage der Schädigungen können Abbildung 4.24b entnommen werden. Die Schädigungen können mit guter Genauigkeit erkannt werden. Nach 14.400 Iterationsschritten wird die untere Grenze des Fitneßwertes von 0,1 unterschritten. Die Rechenzeit für diese Simulation beträgt knapp 18 Stunden. Der Verlauf der Fitneßfunktion in Abbildung 4.24a verdeutlicht die Ganzzahloptimierung. Nach 2000 Iterationsschritten wird die richtige Anzahl der Schädigungen erkannt, die durch den starken Abfall der Fitneßfunktion an dieser Stelle gekennzeichnet ist. Anschließend werden die unbekannten Schädigungsparameter mit guter Genauigkeit identifiziert.

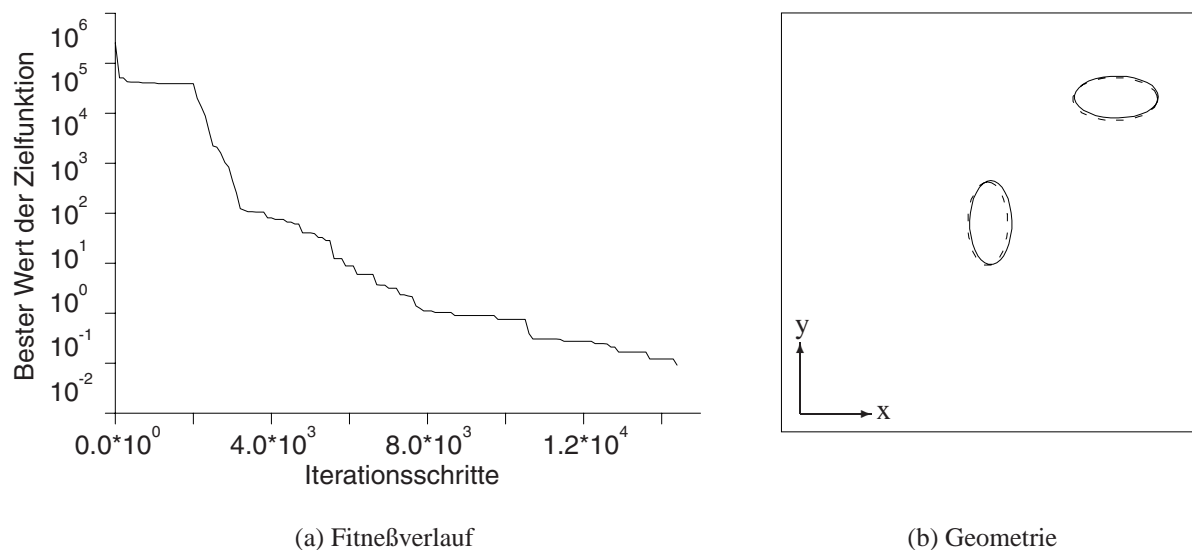


Abbildung 4.24: Identifikationsergebnis von zwei elliptischen Schädigungen bei unbekannter Anzahl an Defekten
 - - - Ergebnis ——— Vorgabe

4.4.2 Identifikation von elliptischen Schädigungen mit einem genetischen Algorithmus im Dreidimensionalen

Nach erfolgreicher Schadenslokalisierung an Hand einer zweidimensionalen Geometrie wird die Identifikation von elliptischen Schädigungen in einem dreidimensionalen Würfel getestet. Die inverse Problemstellung beinhaltet die Erkennung einer Kugel oder eines Ellipsoids innerhalb eines Würfels. Die Schädigungsparameter \mathbf{z}_0 eines Ellipsoids setzen sich aus den drei unbekannten Mittelpunktkoordinaten, den drei Halbachsen und den drei Neigungswinkeln zusammen. In der ersten Simulation beschränken sich die unbekannten Parameter zunächst auf die drei Mittelpunktkoordinaten bei vorgegebenen Radius. Da die Identifikation zum erwünschten Erfolg verlief, wurde die Problemstellung auf die Identifikation des Kugelhalbmessers erweitert. Abschließend wird ein Ellipsoid mit neun Geometriedaten bestehend aus drei Positionsangaben des Mittelpunktes, drei Halbmessern und drei Neigungswinkeln, identifiziert. Es werden bis zu drei Lastfälle zur Minimierung der Fitneßfunktion herangezogen. Die Erfahrung hat gezeigt, daß eine höhere Mutationswahrscheinlichkeit die Konvergenz des genetischen Algorithmus hinsichtlich der Schadenserkenkung verbessern kann. Somit wurde die Mutationswahrscheinlichkeit im Vergleich zu den zweidimensionalen Problemstellungen erhöht.

Eine Fitneßfunktion nach Gleichung 4.3 wird formuliert, die der genetische Algorithmus im Laufe der Optimierung minimieren soll. Der relative Fehler zwischen den gemessenen Verschiebungen $\mathbf{u}_{mes} = \mathbf{u}(\mathbf{z}_0)$ und den vom genetischen Algorithmus angenommenen Verschiebungen \mathbf{u} in Abhängigkeit von den Schädigungsparametern \mathbf{z} ist definiert mit

$$E(\mathbf{z}) = \frac{\sum_{i=1}^{3n} (u_i(\mathbf{z}_0) - u_i(\mathbf{z}))^2}{\sum_{i=1}^{3n} u_i(\mathbf{z}_0)^2}. \quad (4.3)$$

Um den Wertebereich der Fehlerfunktion einzuschränken, wird der quadratische Fehlerwert der Verformungen mit der Summe der Quadrate der Verformungen des Sollzustandes normiert. Auf die Optimierung mit dem genetischen Algorithmus hat diese Normierung keinen Einfluß. Bei der Auswertung der Fitneßfunktion werden an allen n Meßpunkten die Verformungen in allen Koordinatenrichtungen betrachtet, so daß insgesamt $3n$ Meßwerte je Randelementrechnung zur Minimierung der Fitneßfunktion berücksichtigt werden. Bei Anwendung mehrerer Lastfälle zur Schadenserkenkung werden die berechneten Fitneßfunktionen zu einer gesamten Fitneßfunktion aufsummiert.

Die Belastung des Würfels erfolgt nach Abbildung 4.25 zunächst mit nur einem Lastfall. An einer Seite des Würfels wird eine gleichverteilte Zugbelastung aufgebracht, während die gegenüberliegende Seite als fest eingespannt angenommen wird. Da die Optimierung mit dem genetischen Algorithmus eine iterative Vorgehensweise beschreibt und die Randelementrechnung sehr häufig durchgeführt werden muß, wird, um die Rechenzeit ein wenig einzuschränken, eine nicht zu feine aber hinreichende Diskretisierung für die Elementierung des Würfels verwendet.

Der Verlauf der Fitneßfunktion in Abhängigkeit von der Schadensposition wird an Hand eines einfachen Beispiels aufgezeigt. Die Position einer Schädigung wird mit $x_0 = 1,2$, $y_0 = 0,8$ und mit $z_0 = 1,0$ vorgegeben. Die Halbradien des Ellipsoids werden mit $R_1 = 0,2$, $R_2 =$

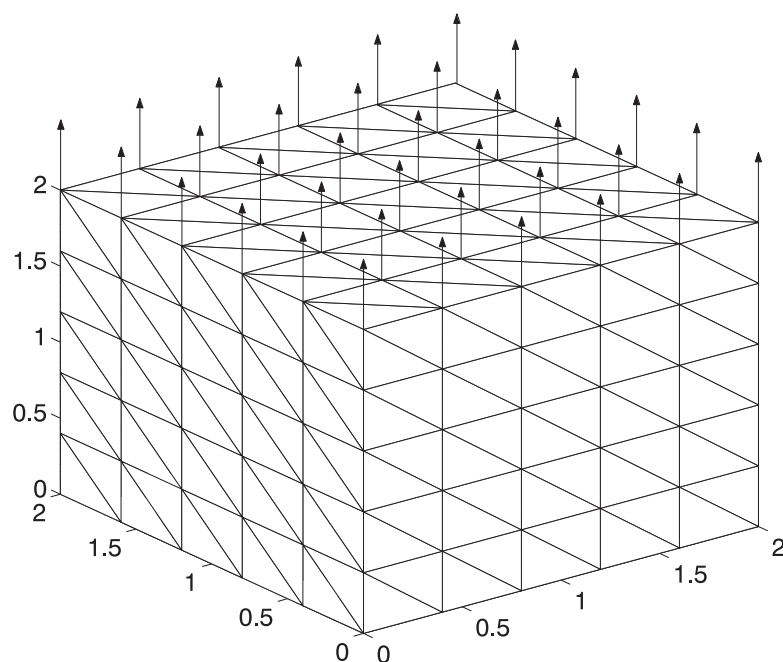


Abbildung 4.25: Belastung und Diskretisierung des Würfels

0,25, $R_3 = 0,25$ festgesetzt, und die Neigung der Schädigung wird zu null angenommen. Abbildung 4.26a zeigt den Verlauf des Fitneßwertes in Abhängigkeit der Schadensposition in x_0 - und y_0 -Richtung. Die anderen charakteristischen Größen des Ellipsoids werden dabei konstant gehalten. In Abbildung 4.26b wird bei der Berechnung des Fitneßwertes nur die z_0 -Koordinate variiert. Beide Abbildungen deuten auf ein konvexes Optimierungsproblem hin, bei dem der Fitneßwert genau ein globales Minimum an der Stelle der zu suchenden Schädigung hat. Bei vorgegebenen richtigen Mittelpunktkoordinaten kann der Fitneßwert nach Abbildung 4.27 in Abhängigkeit zweier Neigungswinkel des Ellipsoids angegeben werden.

Identifikation einer kugelförmigen Fehlstelle mit bekanntem Radius Die einfachste Möglichkeit, eine Fehlstelle in einem Würfel zu lokalisieren, ist die Identifikation eines Ellipsoids, bei der die Radien und die Neigungswinkel vorgegeben sind. Es muß nur noch die Lage der Fehlstelle ermittelt werden. Damit reduziert sich die Anzahl der zu optimierenden Parameter von neun auf drei, die drei Mittelpunktkoordinaten des Ellipsoids. Die Auswirkungen kleinerer Schädigungen im Bauteil auf die Verformungsmeßpunkte ist wesentlich geringer als die größerer Schädigungen. Es kann somit davon ausgegangen werden, daß bei Vorgabe einer größeren Fehlstelle deren Identifikation ermöglicht und beschleunigt werden kann.

Die vom genetischen Algorithmus zu lokalisierende Schädigung hat die Mittelpunktkoordinaten $x_0 = 1,2$, $y_0 = 0,8$ und $z_0 = 1,0$. Die Radien der Fehlstelle werden mit $R_1 = 0,2$, $R_2 = 0,15$ und $R_3 = 0,15$ vorgegeben. Eine Neigung der Hauptachsen der Fehlstelle gegenüber denen des Würfels wird hier nicht betrachtet. Die Mutationswahrscheinlichkeit ist mit zehn Prozent recht hoch gewählt worden. Verschiedene Untersuchungen an diesem Beispiel haben gezeigt, daß eine Erhöhung der Mutationswahrscheinlichkeit zu einer Verringerung der Anzahl

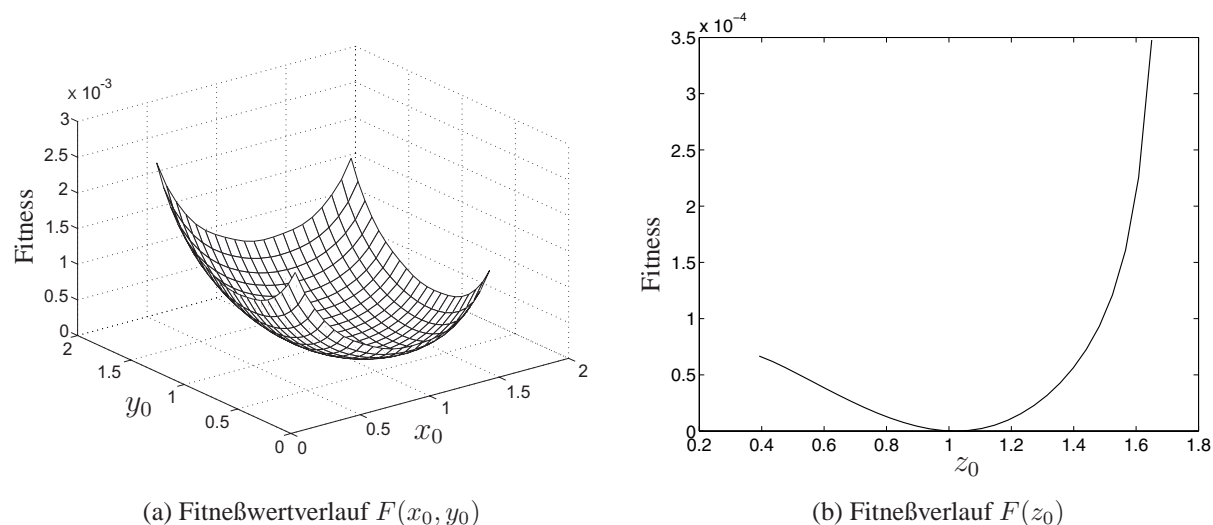


Abbildung 4.26: Fitneßwertverlauf $F(x_0, y_0, z_0)$ in Abhängigkeit von der Schadensposition x_0, y_0, z_0

an Iterationsschritten führt.

Desweiteren wurde der genetische Algorithmus mit unterschiedlicher Anzahl von Individuen getestet. Eine kleinere Individuenanzahl verringert die benötigte Rechenzeit für einen Iterationsschritt. Eine zu kleine Anzahl an Individuen kann aber zum Fehlschlagen der Identifikation oder zu einer erhöhten Anzahl an Iterationsschritten für die Optimierung führen. Abbildung 4.28 zeigt den Verlauf der Identifikation der Fehlstelle innerhalb des in Abbildung 4.25 dargestellten Würfels in Abhängigkeit von der Populationsgröße. In Abbildung 4.28a ist der Verlauf der Fitneßfunktion in Abhängigkeit der Anzahl der Iterationsschritte dargestellt, wobei wie auch in den anderen Darstellungen die dicke durchgezogene Linie die gesuchte Größe angibt. Eine Populationsgröße von $n = 30$ Individuen führt kaum zu einer Verringerung des Fitneßwertes im Laufe von 150 Iterationsschritten. Bei Verwendung einer größeren Anzahl an Individuen verbessert sich der Fitneßverlauf erheblich. Die Fitneßfunktion, simuliert mit einer Populationsgröße von $n = 100$ Individuen, ähnelt bis zum 50. Iterationsschritt dem Verlauf der Fitneßfunktion, die mit $n = 250$ Individuen erstellt worden ist. Erst beim 80. Iterationsschritt kann der Fitneßwert bei 250 Individuen noch einmal deutlich reduziert werden. Dieser niedrige Wert hat aber kaum noch Auswirkungen auf die Genauigkeit der Ergebnisse. Die Abbildungen 4.28b-d zeigen den Verlauf der Mittelpunktkoordinaten bei fortschreitender Anzahl an benötigten Iterationsschritten. Der genetische Algorithmus mit $n = 30$ Individuen kann die y -Koordinate der Fehlstelle nach 150 Iterationen nicht erkennen. Erst bei Verwendung einer Populationsgröße von $n = 100$ wird die Fehlstelle lokalisiert. Es gilt also, die optimale Populationsgröße in Abhängigkeit der Aufgabenstellung zu finden, um einen guten Kompromiß zwischen Rechenzeit und Genauigkeit der Ergebnisse zu erzielen.

Abbildung 4.29 verdeutlicht das Prinzip der genetischen Optimierung. In der linken Spalte ist der Optimierungsverlauf mit einer Populationsgröße von $n = 250$ und in der rechten Spalte

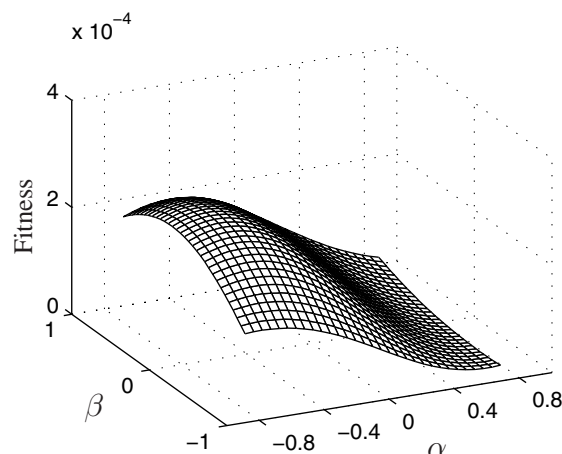


Abbildung 4.27: Fitneßwertverlauf $F(\alpha, \beta)$ in Abhängigkeit von der Neigung der Schädigung

die gleiche Identifikation mit $n = 100$ Individuen dargestellt. Abbildung 4.29a und 4.29b zeigt die Startpopulation des genetischen Algorithmus bevor die ersten genetischen Operatoren angewandt worden sind. Die eingezeichneten Punkte stellen die vom genetischen Algorithmus zu prognostizierenden Mittelpunkte der Schädigung dar. Nach 10 Iterationsschritten ist eine leichte Konzentrierung der Individuen um die zu lokalisierende Schädigung herum zu erkennen. Nach Abschluß der 30. Generation ist die Streuung der einzelnen Individuen schon sehr gering, und die Identifikation verläuft erfolgreich. Entscheidend für die Qualität des Identifikationsergebnisses ist allerdings das Individuum mit dem geringsten Fitneßwert der letzten Population.

Identifikation eines Ellipsoids mit unbekannten Radien Die Optimierung wird nun auf sechs unbekannte Parameter erweitert. Im einzelnen sind wie bisher die drei Mittelpunktkoordinaten einer elliptischen Schädigung und zusätzlich die drei Radien der Fehlstelle zu bestimmen. Der Erfahrung nach ist der Einfluß der Radien auf die Fehlerfunktion nach Gleichung (4.3) geringer als die Veränderung der Lage der Fehlstelle. Die Ausgangsparameter der zu suchenden Fehlstelle sind entsprechend denen der vorherigen Rechnung identisch angenommen zu $x_0 = 1, 2$, $y_0 = 0, 8$ und $z_0 = 1, 0$. Zusätzlich sollen nun die Radien $R_1 = 0, 2$, $R_2 = 0, 25$ und $R_3 = 0, 25$ identifiziert werden. Für den Identifikationsvorgang wird weiterhin nur ein Lastfall verwendet, um einen Vergleich mit der vorherigen Simulation zu ermöglichen.

Die Ergebnisse der Identifikation können den Abbildungen in 4.30 entnommen werden. Abbildung 4.30a zeigt den Fitneßwertverlauf nach Gleichung (4.3). Die Optimierung verläuft erfolgreich. Nach Abschluß der 250 Iterationen kann der Fitneßwert gegenüber dem Anfangswert deutlich reduziert werden. Der Verlauf der Werte der Mittelpunktkoordinaten in Abhängigkeit der Iterationsschritte ist in Abbildung 4.30b gezeigt. Nach Abschluß der genetischen Optimierung ist der Mittelpunkt der elliptischen Fehlstelle lokalisiert worden. Abbildung 4.30c zeigt die erfolgreiche Identifikation der Radien, die in der Belastungsebene liegen. Der Radius, der die Fehlstelle in Zugrichtung beschreibt, kann nach Abschluß der Identifikation nur mit ungenü-

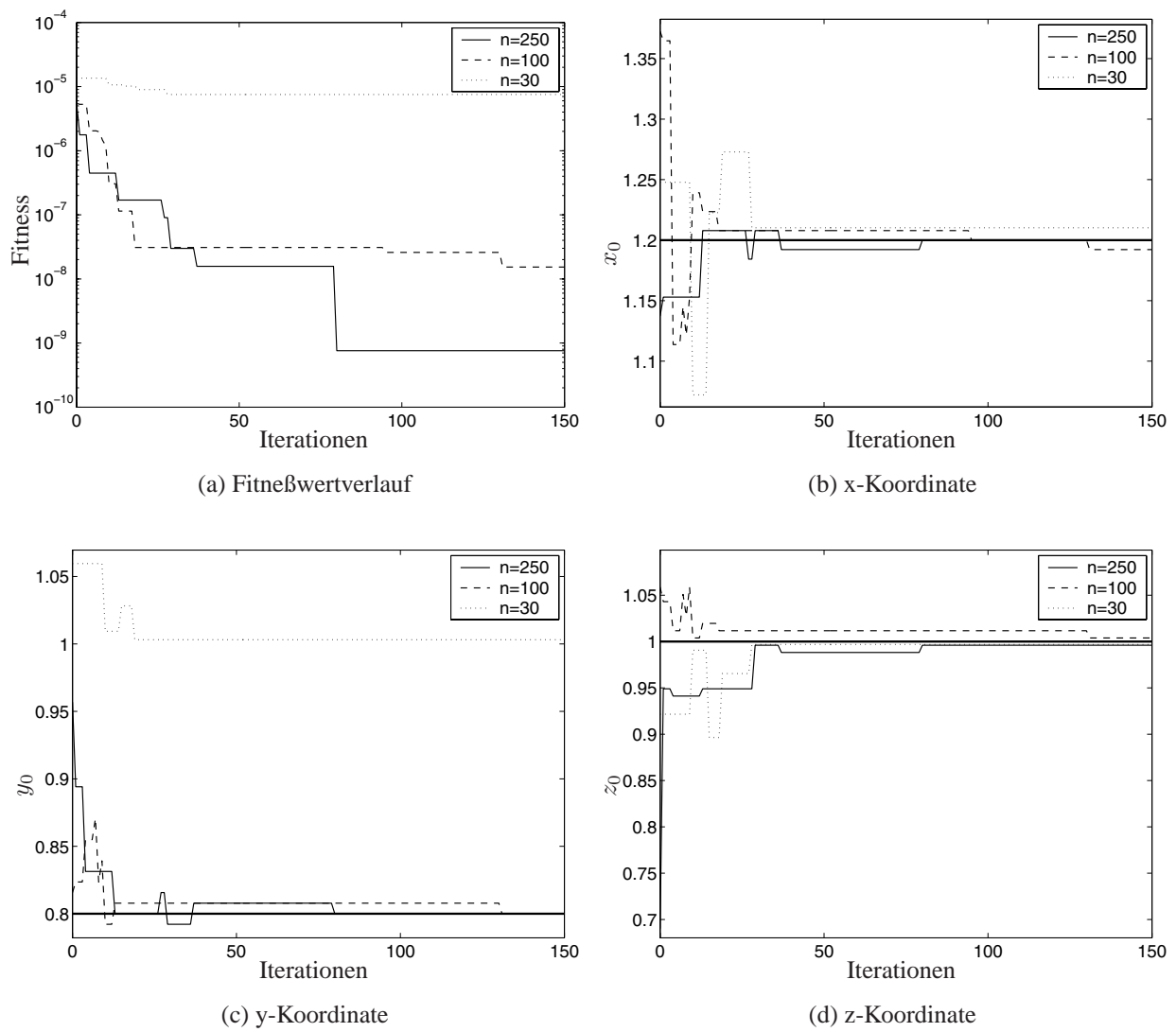


Abbildung 4.28: Ergebnisse der Identifizierung einer Kugel in Abhängigkeit der Populationsgröße

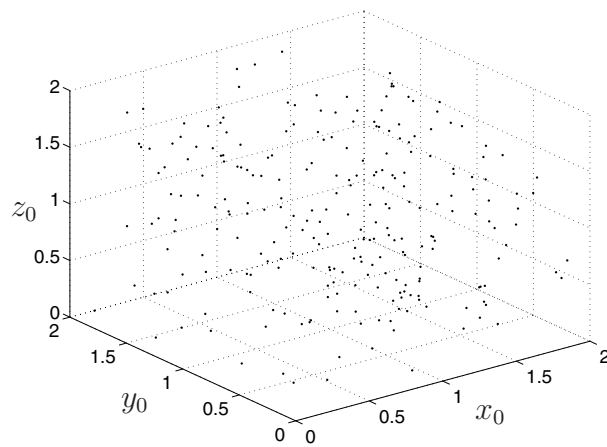
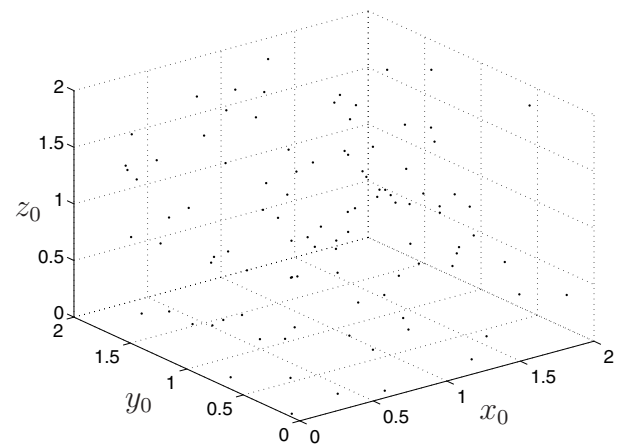
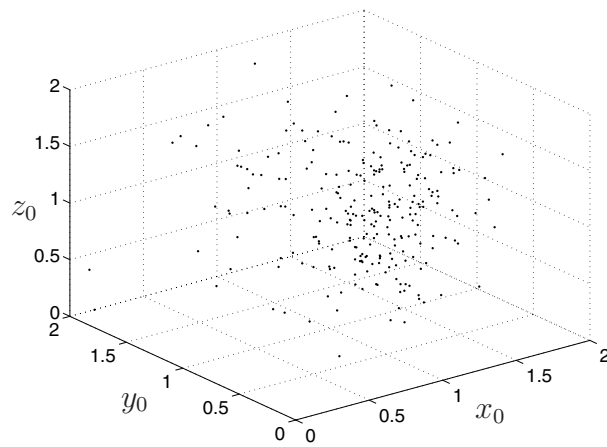
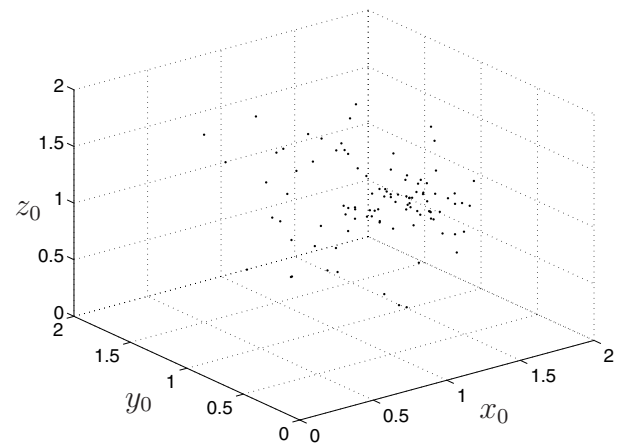
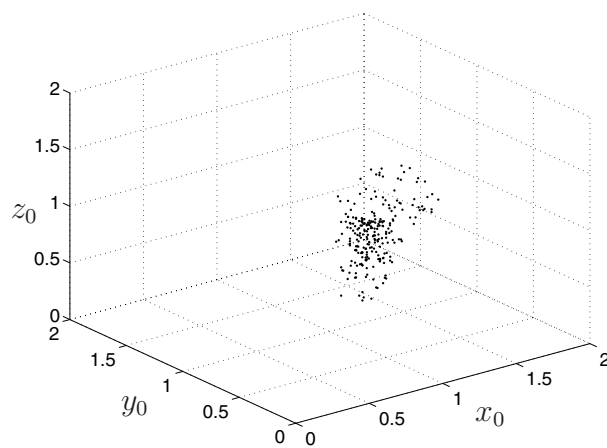
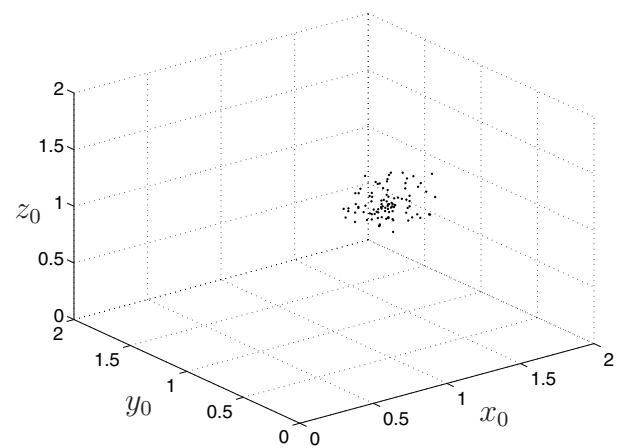
(a) 1. Generation mit $n = 250$ (b) 1. Generation mit $n = 100$ (c) 10. Generation mit $n = 250$ (d) 10. Generation mit $n = 100$ (e) 30. Generation mit $n = 250$ (f) 30. Generation mit $n = 100$

Abbildung 4.29: Verlauf der genetischen Optimierung bei unterschiedlicher Populationsgröße

gender Genauigkeit ermittelt werden. Auffällig ist die schnelle Identifikation der Mittelpunktkoordinaten, deren Werte schon nach knapp 30 Iterationen mit guter Genauigkeit approximiert sind, während die Prognose der Radien erheblich mehr Rechenzeit benötigt. Dieser Tatbestand verdeutlicht den geringeren Einfluß der Radien auf die Fitneßfunktion nach Gleichung (4.3) gegenüber dem Einfluß der Mittelpunktkoordinaten.

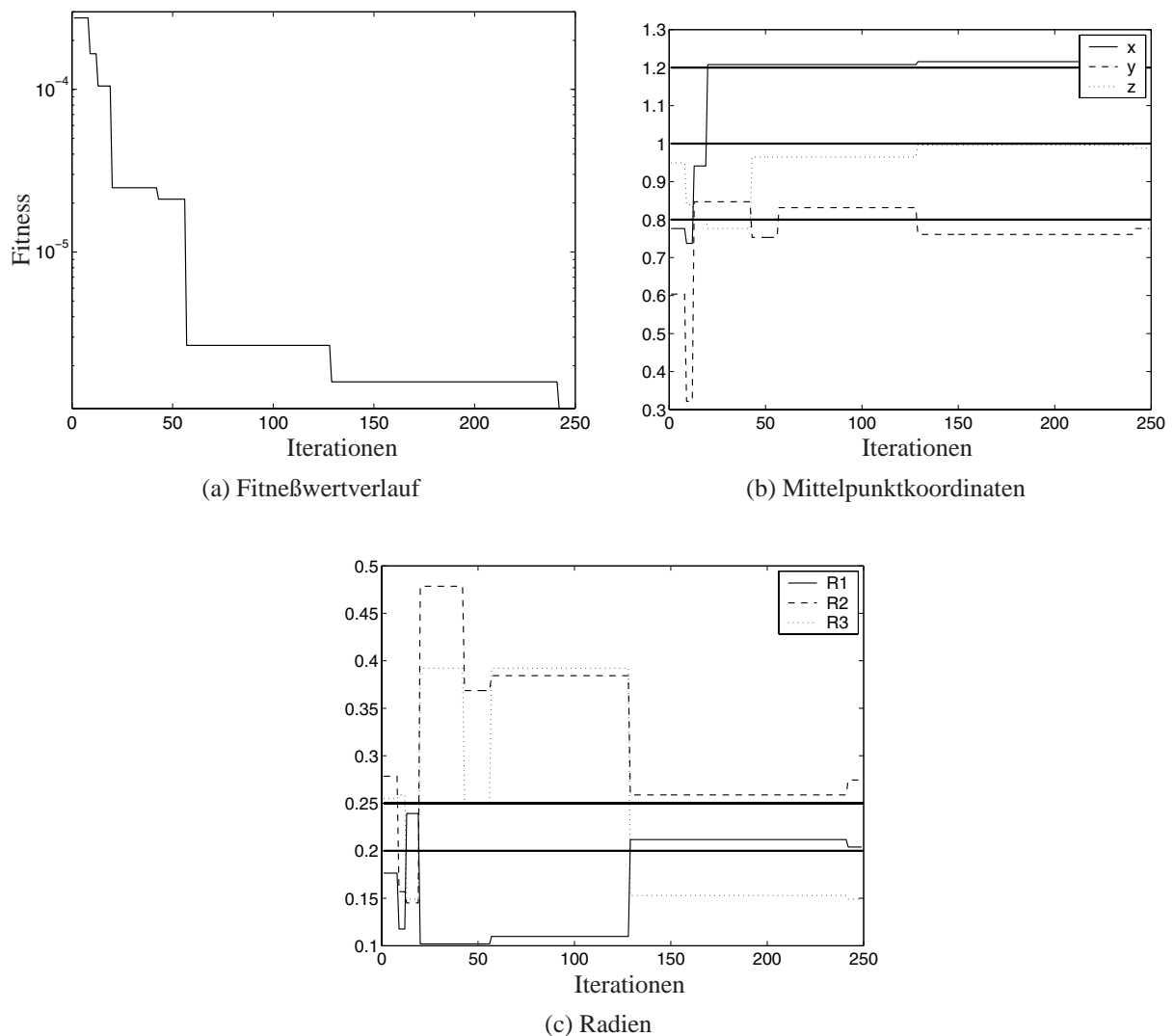


Abbildung 4.30: Identifikation einer elliptischen Fehlstelle mit unbekannten Mittelpunktkoordinaten und Radien

Identifikation eines Ellipsoids mit unbekannten Halbmessern und unbekannter Lage der Hauptachsen Die Lokalisierung einer beliebigen elliptischen Fehlstelle innerhalb eines Würfels beinhaltet die Identifikation der Mittelpunktkoordinaten des Defektes, dessen radiale Aus-

dehnung und dessen Neigung zum Hauptachsensystem des Würfels. Somit müssen neun unbekannte Parameter mit der genetischen Optimierung bestimmt werden. Die Identifikation wird mit drei unterschiedlichen Lastfällen durchgeführt, um ausreichende Informationen über die unbekannten Größen zu erhalten. Der erste Lastfall entspricht der Zugbelastung des Würfels nach Abbildung 4.25. Der zweite und dritte Lastfall beinhaltet die gleiche Belastungsart und -höhe nach Abbildung 4.25. Sie greift allerdings an den anderen beiden Würfeloberflächen an, deren Normalen in die positive Achsenrichtung zeigen. Die jeweils gegenüberliegende Seite wird als fest eingespannt angenommen.

Nach Durchführung mehrerer Tests ist eine optimale Populationsgröße von $n = 1.000$ Individuen für den genetischen Algorithmus vielversprechend. Die Parameter des zu suchenden Ellipsoids sind die Mittelpunktkoordinaten $x_0 = y_0 = z_0 = 1, 2$, die Radien $R_1 = 0, 3$, $R_2 = 0, 2$, $R_3 = 0, 15$ und die Neigungswinkel $\alpha = \beta = 0, 0$, $\gamma = -\pi/6$. Abbildung 4.31 gibt den Verlauf der genetischen Optimierung wieder. Der Fitneßwert nach Abbildung 4.31 kann gegenüber dem Anfangswert zwar deutlich reduziert werden, die Daten der Geometrieparameter des Ellipsoids hingegen werden zum Teil nur mit geringer Genauigkeit erkannt. Der Iterationsverlauf der Mittelpunktkoordinaten ist in Abbildung 4.31b skizziert. Die Lage der Schädigung, die den größten Einfluß auf den Fitneßwert hat, wird nach knapp 40 Iterationen sehr gut prognostiziert. Die Radien werden nach Abbildung 4.31c vom genetischen Algorithmus nicht gut berechnet. Nur der kleinste Radius von $R_3 = 0, 15$ wird mit sehr geringer Abweichung ermittelt. Die Entwicklung der Neigungswinkel in Abhängigkeit von der Anzahl der Iterationsschritte ist in Abbildung 4.31 festgehalten. Die Neigungswinkel des Ellipsoids zum Koordinatensystem des Würfels werden sehr gut erkannt.

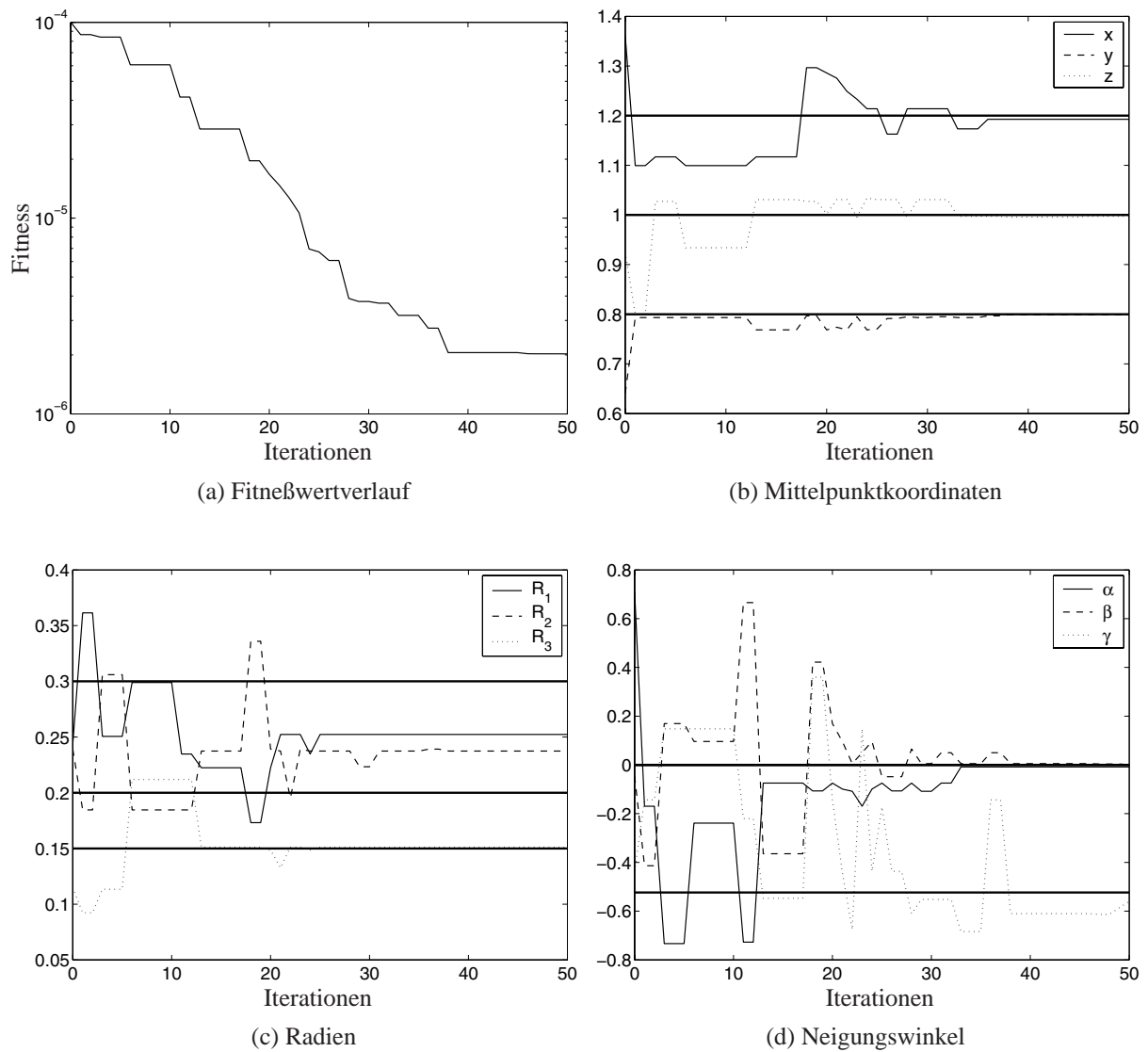


Abbildung 4.31: Identifikation eines Ellipsoids mit unbekannten Mittelpunktkoordinaten, Radien und unbekannter Neigung

Identifizierung einer Fehlstelle in einem L-Profil In einer anderen Problemstellung soll nun gezeigt werden, inwieweit eine Schadenslokalisierung in einer L-förmigen Geometrie möglich ist. Das Bauteil wird nach Abbildung 4.32a durch eine Zugbelastung mit einer kleinen Angriffs-

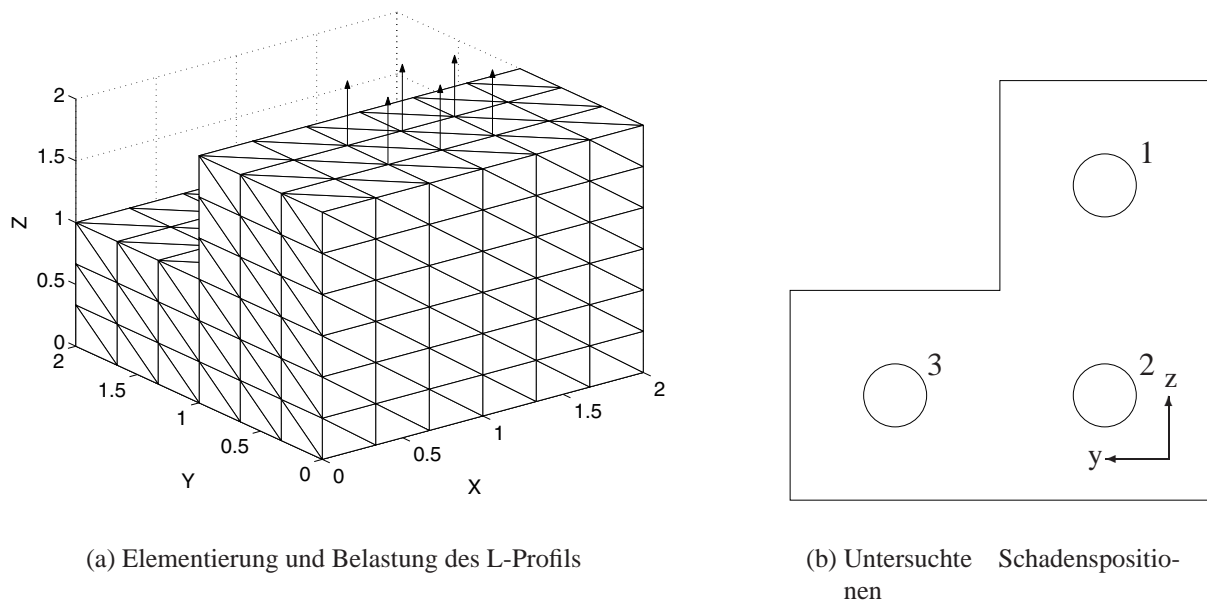


Abbildung 4.32: L-förmige Geometrie: Diskretisierung und Schadensposition

fläche an der Oberseite des Probenkörpers belastet. Auf der gegenüberliegenden Seite $z = 0$ wird der Körper fest eingespannt. Eine nicht zu feine Diskretisierung wird für die Elementierung des Bauteils verwendet, um in der anschließenden inversen Problemstellung zu hohe Rechenzeiten zu vermeiden.

Bei der Untersuchung sollen unterschiedlich vorgegebene Schadenspositionen nach Abbildung 4.32b untersucht werden. Eine Schädigung, die direkt unterhalb der Belastungsstelle liegt, wie in Abbildung 4.32b mit der Positionsnummer 1 gekennzeichnet, sollte leicht zu lokalisieren sein, da diese direkt dem Einfluß der Belastung ausgesetzt ist. Die Identifikation der Schädigung mit der Positionsnummer 3, sollte weniger erfolgreich sein, da diese im Vergleich zu der ungeschädigten Geometrie keine unmittelbaren großen Verformungsänderungen des Randes hervorruft. Weiter werden eine unterschiedliche Anzahl an Meßpunkten betrachtet. Zum einen wird der Verformungswert an jedem Diskretisierungspunkt zur Identifizierung der Fehlstelle herangezogen, zum anderen werden nur die Meßpunkte betrachtet, an denen die Belastung aufgebracht wird.

Zunächst werden die vier Parameter eines Loches, die drei Mittelpunktkoordinaten und der Radius, als die zu identifizierenden Größen vorgegeben.

In Abbildung 4.33a ist das Identifikationsergebnis für die vorgegebene Schädigung in Position 1 dargestellt. Deutlich zu erkennen ist die gute Übereinstimmung der gesuchten Fehlstelle mit der vom genetischen Algorithmus prognostizierten. Für dieses Simulationsergebnis benötigt der

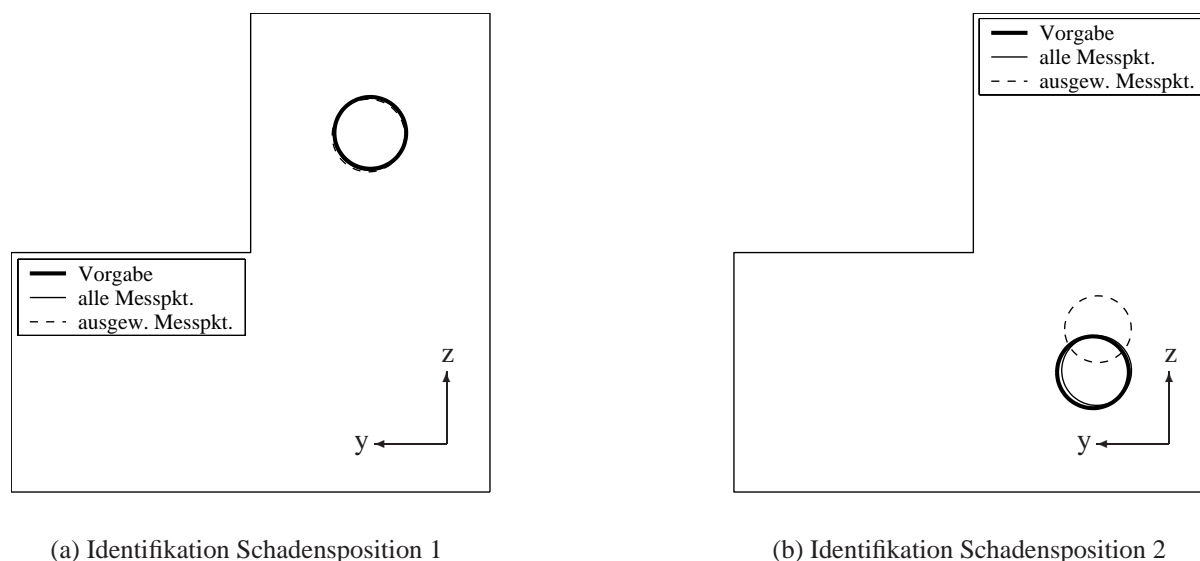


Abbildung 4.33: Identifikation einer Kugel im L-Profil mit unterschiedlicher Anzahl an Meßpunkten

Algorithmus 53 Iterationen unter der Berücksichtigung aller Meßpunkte und 134 Iterationen bei Berechnung mit der eingeschränkten Anzahl an Meßpunkten. An dieser Stelle wird deutlich, daß der Optimierungsvorgang bei Verwendung der eingeschränkten Anzahl an Meßpunkten auf der Oberfläche der Belastung zu höheren Rechenzeiten führen wird, da eine größere Anzahl an Iterationen innerhalb des Optimierungsvorgangs nötig ist.

In Abbildung 4.33b ist das Identifikationsergebnis für die vorgegebene Schädigung in Position 2 dargestellt. Die geringe Anzahl an Meßpunkten auf der Oberfläche reicht nicht aus, um die geometrischen Koordinaten der Kugel mit dem genetischen Algorithmus hinreichend exakt bestimmen zu können. Eine Abweichung der Prognose der Mittelpunktkoordinate in z-Richtung ist zu erkennen. Offensichtlich kann bei der Messung der Verschiebungen nur an der belasteten Oberseite des L-Profils nur ungenügend genau auf die Tiefe der Fehlstelle geschlossen werden. Die y-Mittelpunktkoordinate des Loches wird jedoch sehr gut erkannt.

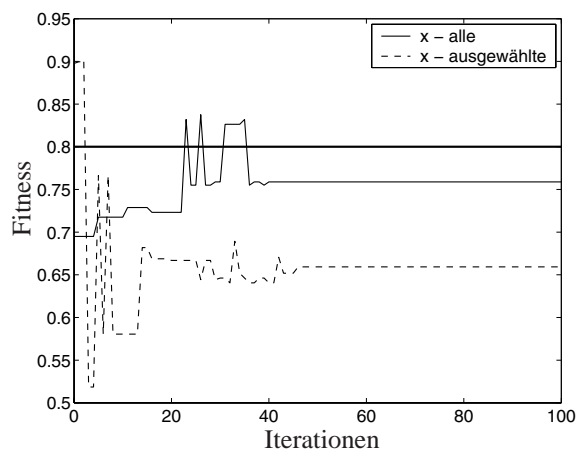
Die Lokalisierung der in Position 3 befindlichen Fehlstelle verläuft weniger erfolgreich, wenn nur die Verschiebungen auf der belasteten Oberfläche verwendet werden. In Abbildung 4.34a-d ist der Optimierungsverlauf der unbekannten Größen dargestellt. Abbildung 4.34a zeigt, daß die x-Koordinate der zu suchenden Schädigung bei Verwendung aller Meßpunkte zur Erstellung der Fitneßfunktion eine bessere Prognose liefert, als mit der Verwendung einer reduzierten Anzahl an Meßpunkten. Der Fehler der Identifikation beläuft sich mit ca. neun Prozent im tolerablen Bereich. Auffällig ist die Abweichung bei der Prognose der Größe der Schädigung. Mit einem Endwert des Kugelradius von $R1 = R = 0,095$ wird die Fehlstelle vom genetischen Algorithmus als wesentlich zu klein erkannt. Dies läßt darauf schließen, daß der Einfluß der Lochgröße auf die Verformungen der belasteten Oberfläche vernachlässigt werden kann.

Bei Erweiterung der Problemstellung auf die Identifikation aller drei Radien eines Ellipsoids, werden alle drei Radien sehr schlecht approximiert. Dieses Ergebnis entspricht im allgemeinen den Erwartungen des Optimierungsverlaufes. Das Ergebnis solch einer Identifikation ist in der Abbildung 4.35 dargestellt. Durch diese Untersuchungen wird eine entscheidende Schwäche des genetischen Algorithmus deutlich. Der Einfluß der Radien auf die Fitneßfunktion ist geringer als die der Mittelpunktkoordinaten der Schädigung. So approximiert der genetische Algorithmus die Mittelpunktkoordinaten der Schädigung mit gutem Resultat aber nicht deren Ausdehnung. Es ist allgemein bekannt und mit diesen Untersuchungen bestätigt, daß der genetische Algorithmus einen hohen Rechenaufwand benötigt, bis auch die Parameter identifiziert sind, die einen geringeren Einfluß auf die Fitneßfunktion haben. Die erfolgreiche Optimierung dieser Fehlerfunktion bis zu einer ganz geringen Abweichung des Optimums vom Zielwert ist mit der genetischen Strategie nahezu auszuschließen. Die bis hierhin gewonnenen Ergebnisse verdeutlichen den Einfluß der Anzahl gleichverteilter Meßpunkte am Rand des Prüfkörpers zur genauen Identifizierung der Fehlstelle.

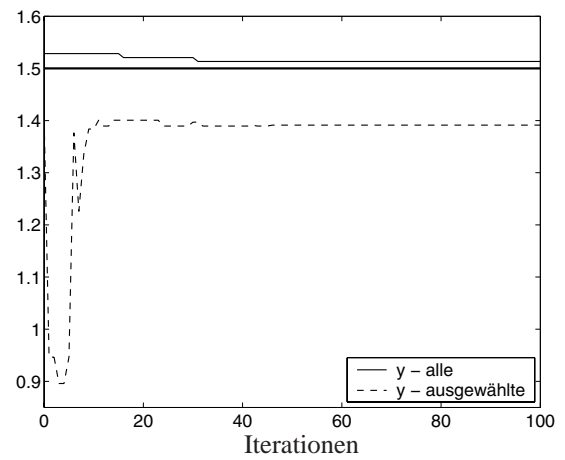
Es wird nun die Auswirkung eines weiteren Belastungsfalls auf die Güte des Optimierungsverlaufes durch den genetischen Algorithmus untersucht. Die zusätzliche Zugbelastung erfolgt nach Abbildung 4.32 an der anderen schmalen Oberfläche L-Profils, die durch Punkte des Profils mit einer y -Koordinate von $y = 2$ definiert ist. Zur Berechnung der Fitneßfunktion werden alle Diskretisierungspunkte als Meßpunkte herangezogen. Durch diese zusätzlichen Informationen kann die Identifikation einer Fehlstelle an Schadensposition 3 mit gutem Resultat abgeschlossen werden. Die prognostizierte Lage der Kugel kann Abbildung 4.36 entnommen werden.

Zusätzlich zu den vorherigen zwei Lastfällen kann die Schadenserkenkung mit einem weiteren Lastfall durchgeführt werden. Die Zugbelastung greift an der Stirnseite des L-Profils an. Abbildung 4.37 stellt das Identifikationsergebnis eines Ellipsoids an der Schadensposition 3 dar.

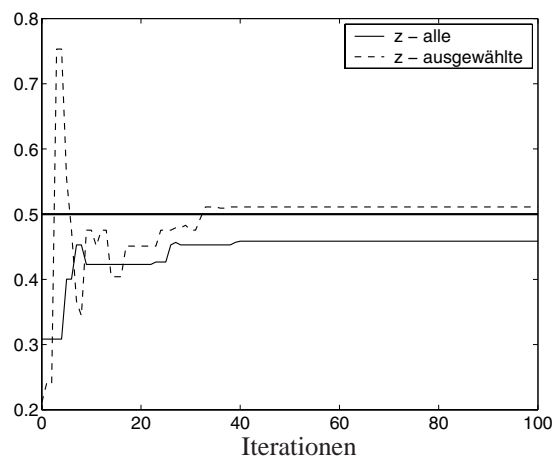
Die Genauigkeit der Ergebnisse kann gegenüber Abbildung 4.35 deutlich verbessert werden. Die Fehler der Prognose zweier der drei Radien ist aber weithin recht hoch.



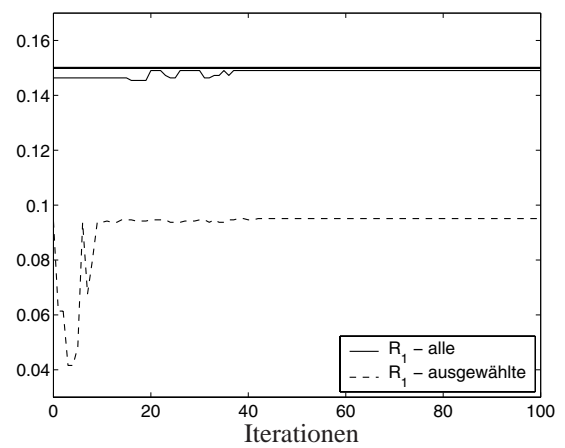
(a) x-Koordinate



(b) y-Koordinate



(c) z-Koordinate



(d) Radius

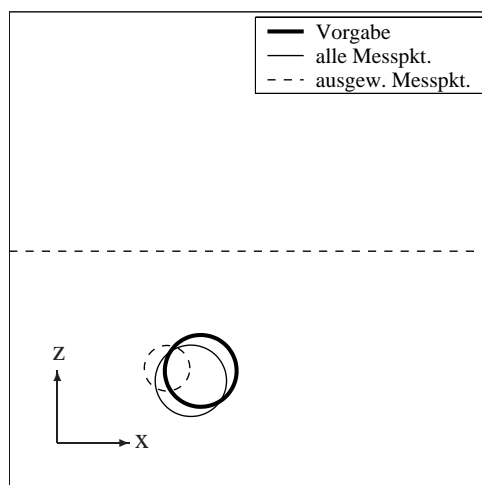
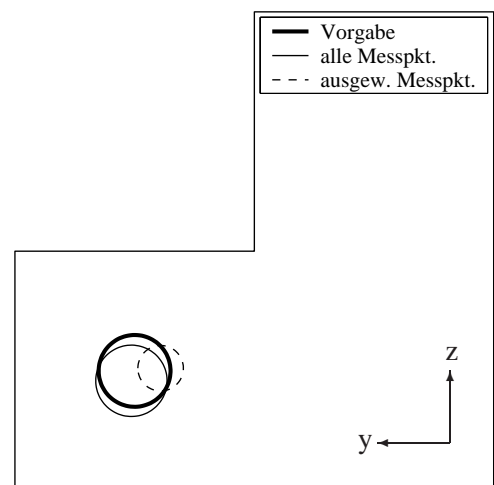
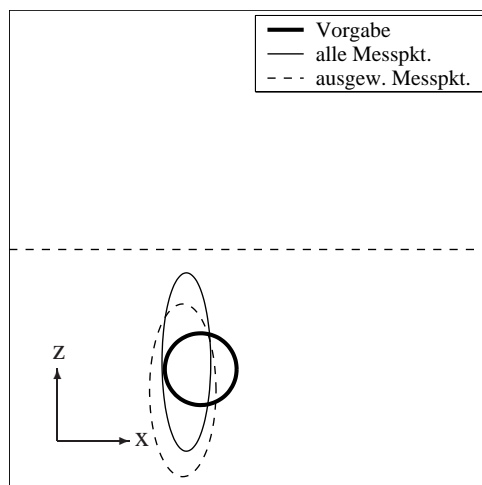
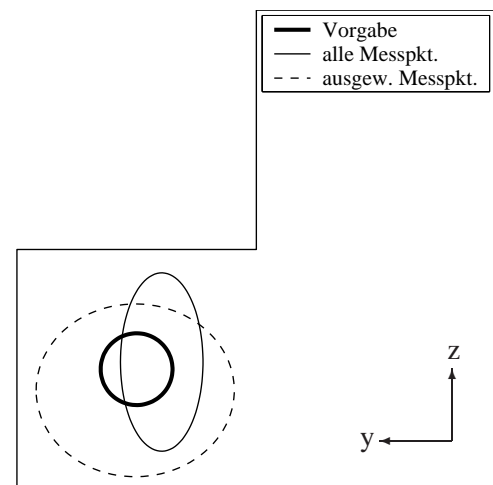

(e) Identifikation Schadensposition 3
Vorderansicht

(f) Identifikation Schadensposition 3 Sei-
tenansicht

Abbildung 4.34: Identifikation einer Kugel im L-Profil mit unterschiedlicher Anzahl an Meß-
punkten

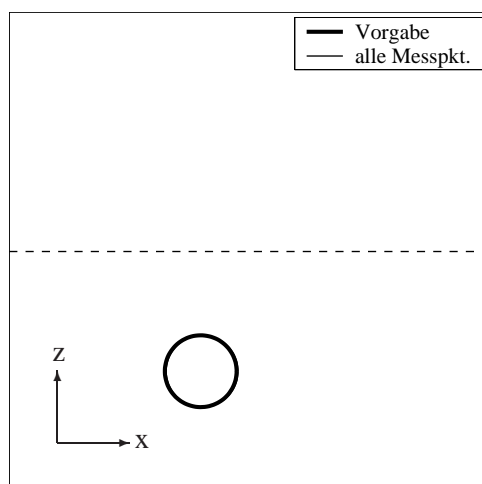


(a) Identifikation Schadensposition 3
Vorderansicht

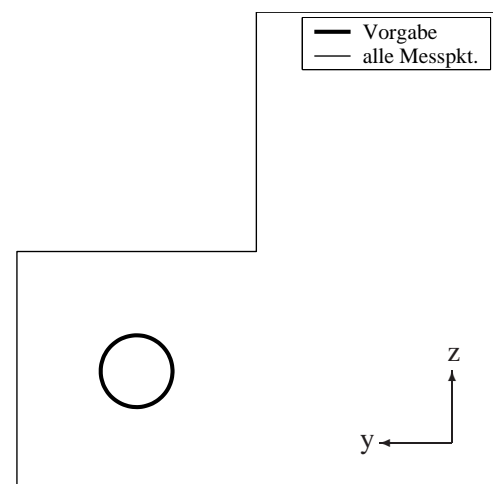


(b) Identifikation Schadensposition 3
Seitenansicht

Abbildung 4.35: Identifikation eines Ellipsoids im L-Profil mit unterschiedlicher Anzahl an Meßpunkten

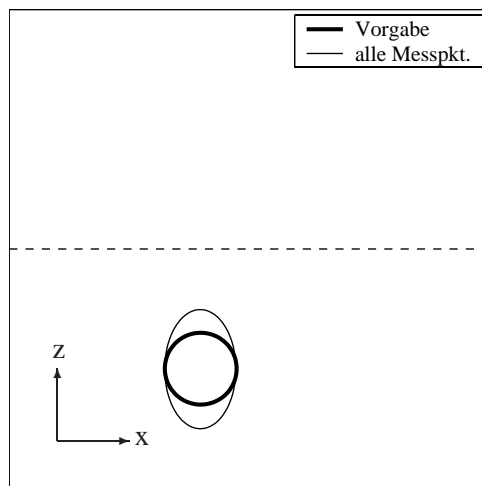


(a) Identifikation Schadensposition 3
Vorderansicht

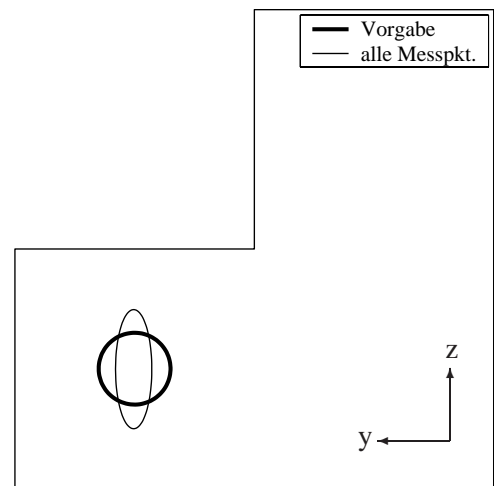


(b) Identifikation Schadensposition 3
Seitenansicht

Abbildung 4.36: Identifikation eines Ellipsoids im L-Profil unter Verwendung zweier Lastfälle und aller Meßpunkte



(a) Identifikation Schadensposition 3
Vorderansicht



(b) Identifikation Schadensposition 3
Seitenansicht

Abbildung 4.37: Identifikation eines Ellipsoids im L-Profil unter Verwendung dreier Lastfälle und aller Meßpunkte

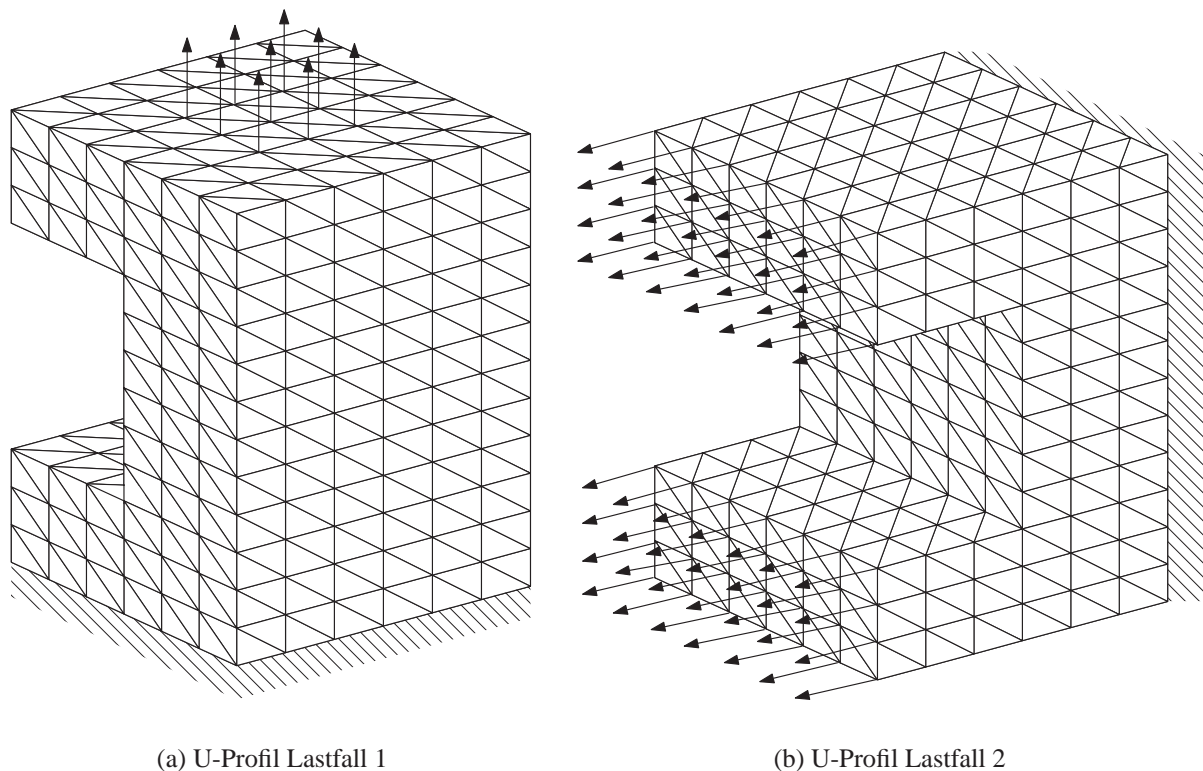


Abbildung 4.38: Diskretisierung und Belastung des U-Profils

Identifikation einer Fehlstelle in einem U-Profil In einer weiteren Untersuchungsreihe soll die Schädigungsidentifizierung an einer U-förmigen Geometrie durchgeführt werden. Ein U-Profil stellt eine kompliziertere Geometrie als die zuvor betrachtete L-Form dar. Einfache Untersuchungen haben ergeben, daß es mit dem genetischen Algorithmus, ähnlich wie in den vorherigen Untersuchungen am L-Profil, Schwierigkeiten gibt, die Schadensidentifikation erfolgreich durchzuführen. Daher wird zur Erkennung der Schadensparameter eine kombinierte Optimierungsstrategie bestehend aus dem genetischen Algorithmus als globale und dem Kalman-Filter als lokale Optimierungsstrategie eingesetzt. Zunächst wird der genetische Algorithmus zur Identifikation der Schädigung eingesetzt. Somit wird ein hinreichend guter Startwert für die weitere Optimierung mittels des Kalman-Filters gefunden. Das Kalman-Filter wird hier eingesetzt, um die Schädigung mit einer guten Genauigkeit zu lokalisieren.

Die Belastung des Probekörpers erfolgt nach Abbildung 4.38, wobei die der Belastung jeweils gegenüberliegende Seite fest eingespannt wird. Es werden zusammenfassend nur die Untersuchungen aufgeführt, deren Schädigung aus neun zu bestimmenden Parametern bestehen. Dies entspricht der Lokalisierung einer Schädigung mit drei zu bestimmenden Mittelpunktkoordinaten, drei unbekannten Halbachsen und drei Neigungswinkeln. In Abbildung 4.39 sind die Ergebnisse der Identifizierung einer elliptischen Fehlstelle unter Belastung des U-Profils nach Lastfall 1 dargestellt. Die zu suchende Fehlstelle befindet sich, wie in Abbildung 4.40 dar-

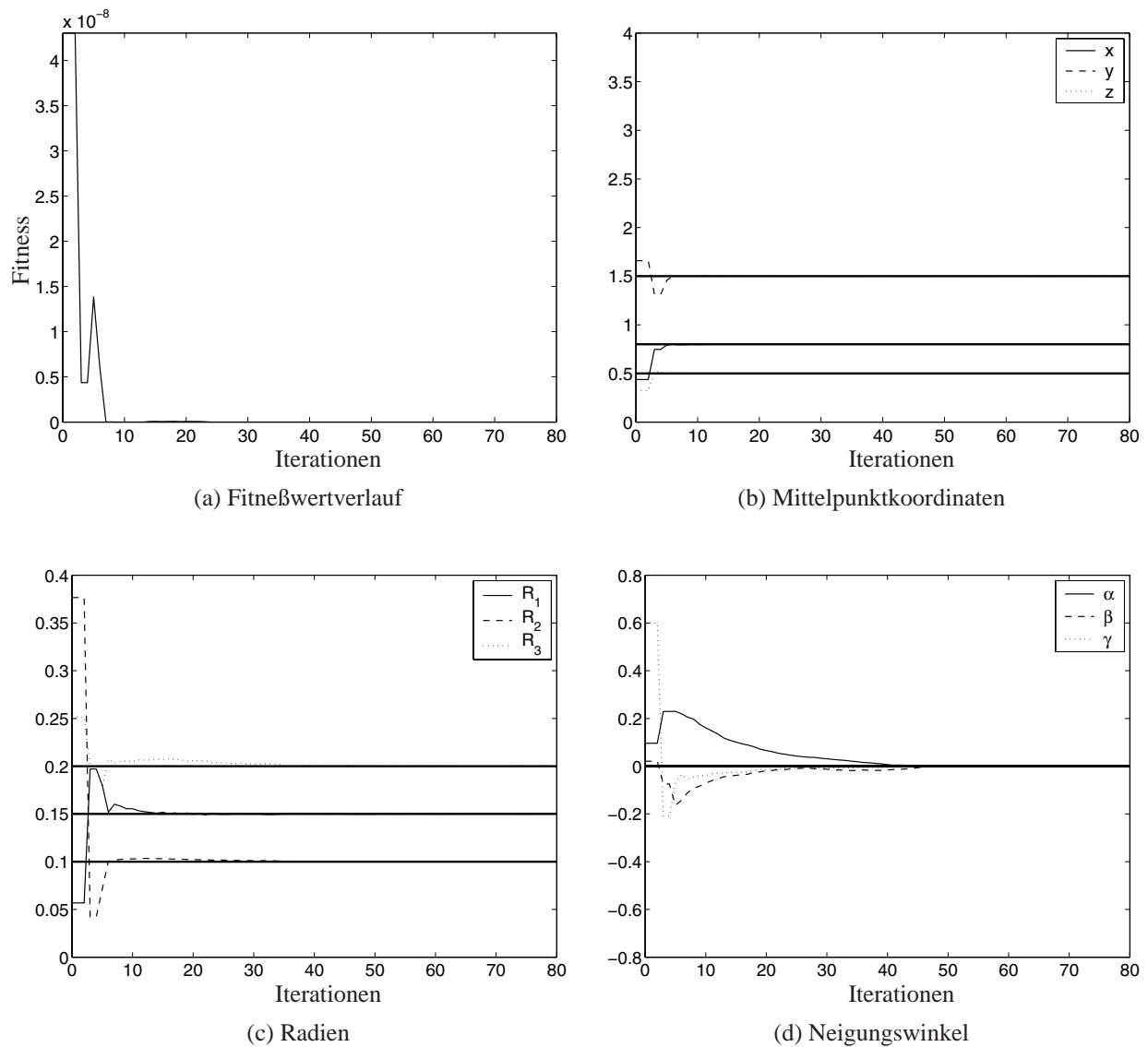


Abbildung 4.39: Identifikation einer elliptischen Fehlstelle im Flansch eines U-Profiles mit unbekannten Mittelpunktkoordinaten, Radien und Neigungswinkeln durch die hybride Optimierungsstrategie

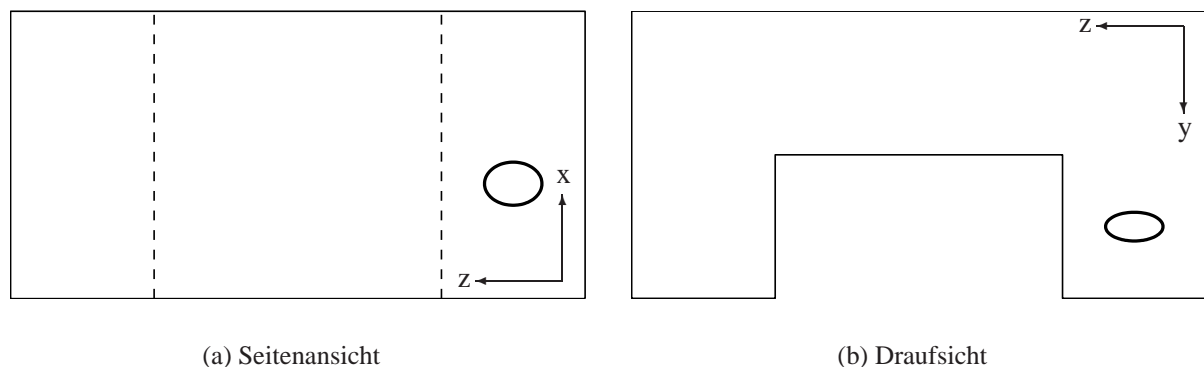


Abbildung 4.40: Ansichten der zu suchenden und gefundenen Fehlstelle innerhalb des U-Profils

gestellt, im nicht belasteten Flansch des U-Profils. Der genetische Algorithmus wird für die ersten zehn Iterationen eingesetzt. Die Wahl eines Startwertes für die Optimierung entfällt. Abbildung 4.39a zeigt die deutliche Minimierung der Fehlerfunktion innerhalb der ersten zehn Iterationen durch den genetischen Algorithmus. Zu diesem Zeitpunkt der Optimierung sind die Mittelpunktkoordinaten, die den größten Einfluß auf den Fehlerwert haben, identifiziert. Dieses Ergebnis entspricht den vorherigen Schadensidentifikationen in einem Würfel. Der nach zehn Iterationen einsetzende Kalman-Filter optimiert nun die Größen der Halbachsen und der Neigungswinkel. Die Größen der Halbachsen werden erst nach ca. 35 Iterationen des Kalman-Filters gut erkannt. In Abbildung 4.39c ist der Optimierungsverlauf der Halbachsen der Fehlstelle dargestellt. Die Zielgrößen werden durch die waagrecht eingezeichneten Linien skizziert. Die Prognose der Neigungswinkel ist in Abbildung 4.39d wiedergegeben. Die Winkel, die den geringsten Einfluß auf den Fehlerwert haben, werden erst nach über 50 Iterationen hinreichend gut durch das Kalman-Filter identifiziert.

In einem zweiten Beispiel sollen die Parameter einer Fehlstelle in dem Steg des U-Profils identifiziert werden. Das U-Profil wird nach Abbildung 4.38b durch Lastfall 2 belastet. Die Identifikation dieser Fehlstelle stellt eine schwierigere Aufgabe dar als die Erkennung einer Fehlstelle innerhalb des Flansches. Ein im Flansch gelegener Defekt wird direkt dem Kraftfluß der Belastung ausgesetzt und hat somit große Auswirkungen auf die Verformungen der Außenseite des Profils. Diese Tatsache gilt nicht für den einen innerhalb des Steges gelegenen Defekt.

Die Einstellungen der hybriden Optimierungsstrategie werden im Vergleich zur vorherigen Simulation nicht geändert. Die ersten zehn Schritte der Optimierung werden mit dem genetischen Algorithmus durchgeführt. Anschließend wird das Kalman-Filter zur genauen Schadenserkennung angewendet. In Abbildung 4.42 ist das Ergebnis der Schadenserkennung abgebildet. Ähnlich der vorherigen Untersuchung ist eine starke Reduzierung des Fehlerwertes nach Abbildung 4.42a innerhalb der ersten Iterationen des genetischen Algorithmus zu erkennen. Die Mittelpunktkoordinaten der Schädigung können nach Abbildung 4.42 genauso schnell gefunden werden wie in der vorherigen Rechnung. So ist die Position bei der Übergabe der Startwerte für den Kalman-Filter sehr gut identifiziert. In den verbleibenden 30 Iterationen werden mittels des Kalman-Filters die Größen der Halbachsen und deren Neigung bestimmt. Abbildung 4.42c zeigt

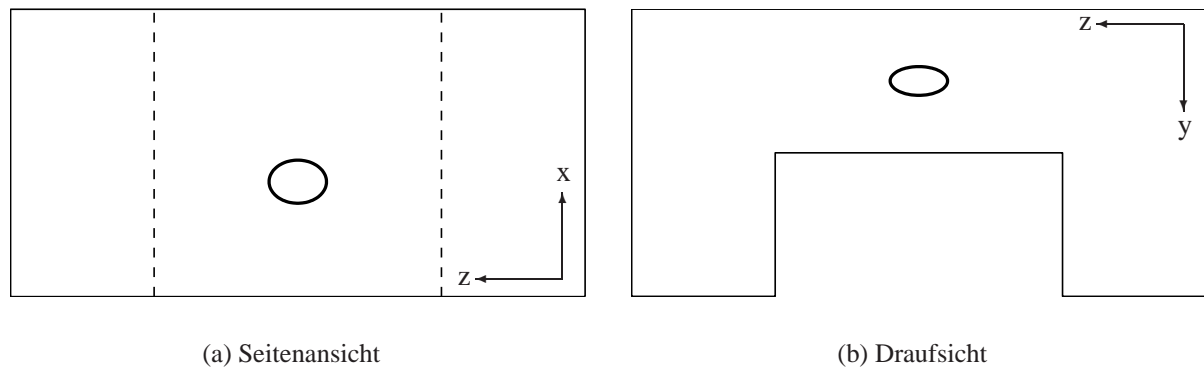
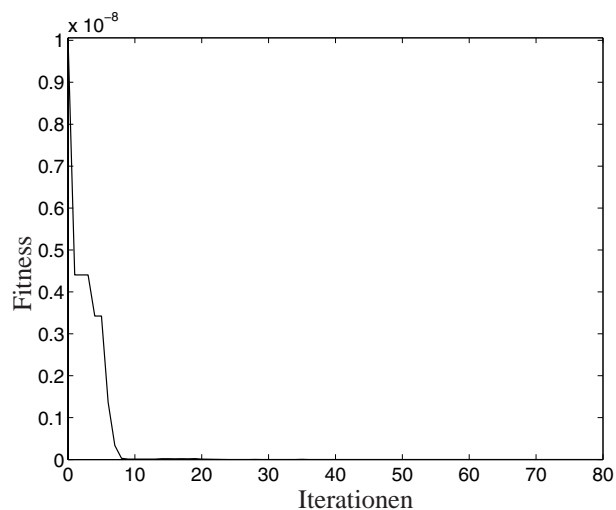
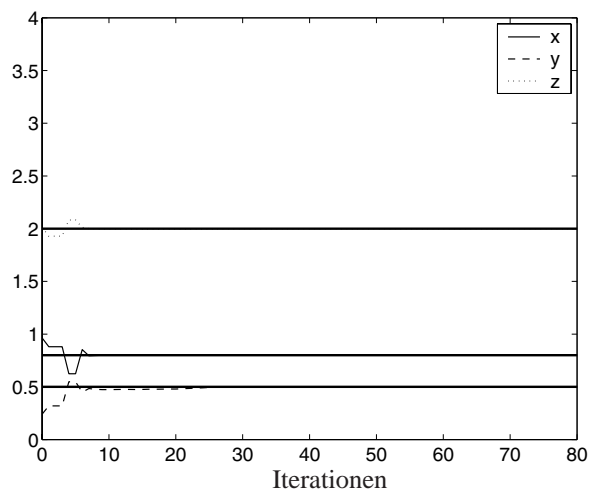


Abbildung 4.41: Ansichten der zu suchenden und gefundenen Fehlstelle innerhalb des U-Profiles

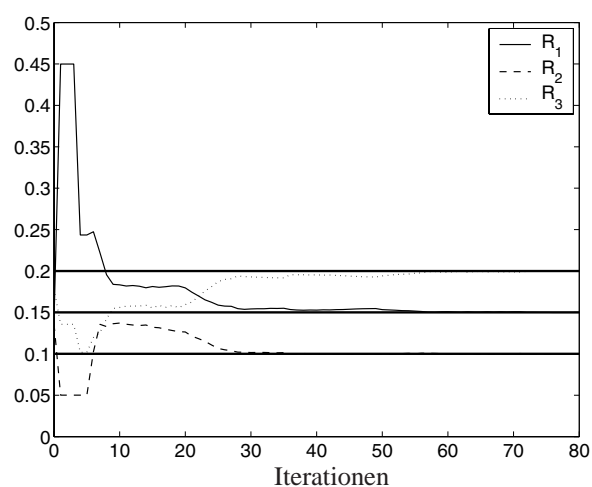
den Optimierungsverlauf der Halbachsen. Diese werden ähnlich dem vorangegangenen Beispiel in kürzerer Zeit identifiziert als die Neigungswinkel nach Abbildung 4.42d. Die angeführten numerischen Berechnungen mit der hybriden Optimierungsstrategie zeigen ein sehr gutes Konvergenzverhalten. Die Anwendung dieser kombinierten Strategie verkürzt zunächst die benötigte Rechenzeit und ermöglicht bei komplizierteren Geometrien überhaupt eine Schadensidentifikation. Beide Verfahren einzeln angewandt würden zu einem Versagen der Schadensidentifikation führen. Der genetische Algorithmus würde bei der Genauigkeit der zu prognostizierenden Parameter zu viel Rechenzeit benötigen, während das Kalman-Filter die Optimierung in einem lokalen Minimum beenden würde.



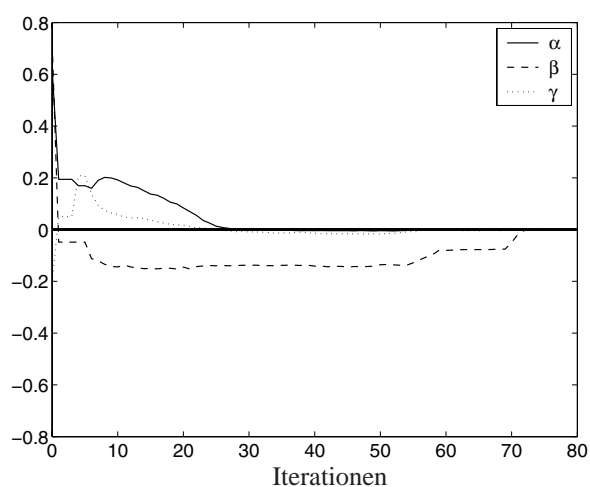
(a) Fitneßwertverlauf



(b) Mittelpunktkoordinaten



(c) Radien



(d) Neigungswinkel

Abbildung 4.42: Identifikation einer elliptischen Fehlstelle im Steg eines U-Profiles mit unbekannten Mittelpunktkoordinaten, Radien und Neigungswinkel durch die hybride Optimierungsstrategie

5 Loch- und Rißerkennung im Zeitbereich

Nach der Schadenserkennung mit statischen Belastungsfällen soll nun die Schadenserkennung mit zeitlich veränderlichen Belastungen untersucht werden. Die Vorgehensweise der Schadenserkennung ist mit der von Kapitel 4 weitgehend identisch.

Für die Durchführung der Optimierung muß ein „Experiment“ berechnet werden. Im Gegensatz zu Kapitel 4 muß eine Randelementsimulation im Zeitbereich durchgeführt werden. Die nun zeitlich abhängigen Verschiebungen werden in u_0 gespeichert. An dieser Stelle wird noch einmal darauf hingewiesen, daß keine Laborexperimente durchgeführt werden. Das „Experiment“ soll zu einem späteren Zeitpunkt durch ein Laborexperiment ersetzt werden. In dieser Arbeit wird es noch mit der Randelementmethode simuliert.

Die Ergebnisse des „Experimentes“ werden zur Durchführung der Optimierung benötigt. Das neuronale Netz nimmt auch hier eine Sonderstellung ein. Es müssen Lerndaten und Testdaten generiert werden, um das Netz zu erstellen und dessen Leistungsfähigkeit zu überprüfen. Ein anschließendes Laborexperiment könnte die Anwendung in der Praxis überprüfen.

Aufgrund der zu erwartenden hohen Rechenzeiten bei Verwendung des genetischen Algorithmus zur Schadenserkennung wird dieser hier nicht weiter betrachtet. Wenn die Rechnerleistung in Zukunft wie erwartet steigen wird, können Simulationen mit parallelisierten genetischen Algorithmen mit vertretbaren Rechenzeiten durchgeführt werden.

Da das SQP-Verfahren in der Statik im Vergleich zum Kalman-Filter eine langsamere Konvergenzgeschwindigkeit aufzeigt, wird es in den folgenden Untersuchungen ebenfalls nicht weiter berücksichtigt.

5.1 Identifikation mit dem linearisierten Kalman-Filter

Das Kalman-Filter nach Kapitel 3.2.2 wird zur Schadenslokalisierung von Rissen innerhalb dynamisch belasteter Scheiben verwendet. Die Untersuchung der Rißidentifikation mittels des Kalman-Filters beinhaltet die Erkennung einer Schädigung oder mehrerer Schädigungen innerhalb einer zweidimensionalen Struktur. Dabei wird unterschieden zwischen einer elliptischen und einer rißförmigen Schädigung. Ein möglicher Kontakt der Rißflächen wird durch den Kontaktalgorithmus nach Kapitel 2.3 berücksichtigt. Es wird der Einfluß unterschiedlicher Belastungsarten auf das Identifikationsergebnis untersucht. Abschließend werden nichtkonvexe Geometrien zur Schadenserkennung gewählt.

Die zu suchenden Parameter sind für einen Riß dessen Mittelpunktkoordinaten, dessen Länge und der Neigungswinkel zur Horizontalen. Eine ellipsenförmige Schädigung wird durch die Lage ihres Mittelpunktes und die Größe der Halbachsen definiert. Die Startwerte für das Kalman-Filter werden bei den numerischen Untersuchungen in eckigen Klammern angegeben. Dabei gilt für den Riß die folgende Notation: $[x_c \ y_c \ l \ \alpha]$. x_c und y_c geben die Mittelpunktkoordinaten an, l die Rißlänge und α den Neigungswinkel des Risses zur x -Achse. Bei der Identifikation mehrerer Risse erweitert sich dementsprechend die Datenmenge. Die Startwerte der Ellipse werden in folgender Notation angegeben: $[x_c \ y_c \ a \ b]$. Der Startwert des Ellipsenmittelpunktes wird mit x_c und y_c angegeben. Die Durchmesser der Ellipse zu Beginn der Optimierung werden durch a und b dargestellt.

Mit der Kenntnis der Verformungen der geschädigten Scheibe aus dem „Experiment“ wird nun mit der Optimierung begonnen. Das mechanische Verhalten der Scheibe wird mit der gleichen wie im Experiment verwendeten Belastung mit der Randelementmethode simuliert. Die daraus gewonnenen Werte der Verschiebungen werden durch das Kalman-Filter mit den experimentell ermittelten Größen verglichen. Dadurch werden während der Berechnung durch das Kalman-Filter die Rißparameter so optimiert, daß am Ende die Schädigungsparameter des „Experimentes“ ermittelt werden können. Als Kennwert für die Qualität eines Schätzwertes wird der Fehlerwert F

$$F = \sum_{k=1}^n \sum_{i=1}^{2m} (u_{ik,\text{Messung}} - u_{ik,\text{Schätzung}})^2 \quad (5.1)$$

benutzt, in dem die Verformungen aus dem Experiment mit $u_{ik,\text{Messung}}$ bezeichnet wurden. Die Verformungen der vom Kalman-Filter prognostizierten Schadenssituation werden mit $u_{ik,\text{Schätzung}}$ gekennzeichnet. Es werden alle m Diskretisierungspunkte in beide Koordinatenrichtungen zu jedem Zeitschritt k zum Auswerten der Fehlerfunktion herangezogen.

5.1.1 Modellierung der Struktur

Die durchgeführten Untersuchungen beschränken sich wie oben erwähnt auf zweidimensionale elastische Scheiben. Als einfachste Form werden die Beispielrechnungen an einer quadratischen Scheibe mit den Abmessungen 100×100 vorgenommen. Im Anschluß werden L-förmige und U-förmige Scheiben betrachtet. Zur Beschreibung des elastischen Verhaltens werden die Materialkennwerte Elastizitätsmodul zu $E = 1 \cdot 10^6$, die Querkontraktionszahl $\nu = 0,3$ und die Dichte $\rho = 2.450$ angenommen.

Die Randelementmethode setzt eine Diskretisierung des Randes und der Zeit voraus. In [2] wird eine Maßzahl β angegeben, durch die die beiden Diskretisierungen in einen mathematischen Zusammenhang gesetzt werden

$$\beta = \frac{c_1 \Delta t}{r_e}. \quad (5.2)$$

In Gleichung (5.2) bezeichnet Δt die äquidistante Zeitschrittweite, c_1 die Geschwindigkeit der durch die dynamische Belastung hervorgerufenen Kompressionswelle und r_e die mittlere Länge der Randelemente. Nimmt β den Wert $\beta = 1,0$ an, so durchläuft die Kompressionswelle in

einem Zeitschritt genau eine Elementlänge. Dieser Wert sollte nicht überschritten werden. Der zeitliche Verlauf der Verschiebungen könnte stark geglättet werden gegenüber der tatsächlichen Lösung, da nicht jeder durch die Wellenreflektionen auftretender Effekt korrekt berücksichtigt werden kann. Ebenso sollten Werte mit $\beta < 0,8$ vermieden werden, da numerische Instabilitäten auftreten können. Für die hier betrachtete Problemstellung haben einige Testrechnungen ergeben, daß der Wert $\beta = 0,9$ sehr gut geeignet ist.

Für den äußeren Rand der quadratischen Scheibe wird eine Diskretisierung mit 20 Punkten pro Seite gewählt. Zur Annäherung des Verschiebungs- und Spannungsverlaufes werden isoparametrische quadratische Elemente verwendet. Folglich wird eine Scheibenseite, wie in Abbildung 5.1 dargestellt, mit zehn Elementen der Länge zehn diskretisiert. Aus Gleichung (5.2) ergibt sich dann für einen Quotienten $\beta = 0,9$ und mit der aus den Materialkennwerten resultierenden Wellengeschwindigkeit $c_1 = 23,44$ die Zeitschrittweite $\Delta t = 0,384$. Die Ränder der Fehlstel-

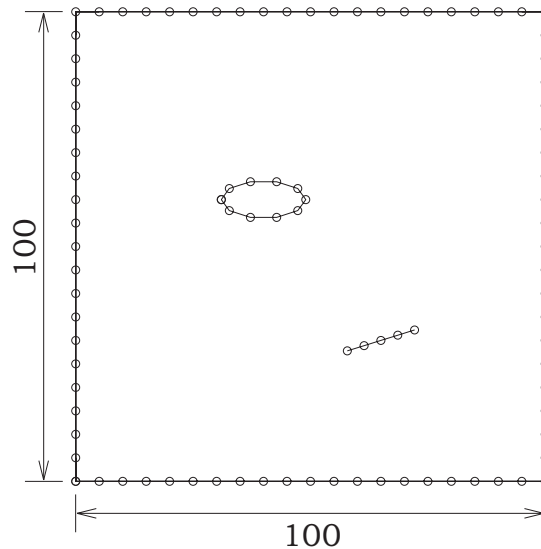


Abbildung 5.1: Diskretisierung des Randes und der Schädigungen

len bestehen für ellipsenförmige Fehlstellen aus fünf und für rißförmige aus zwei quadratischen Elementen. Sowohl für den äußeren Rand als auch für den inneren wurden feinere Diskretisierungen getestet. Die dadurch erzielte genauere Berechnung der Randwerte der Struktur führte allerdings zu keiner Verbesserung der Schädigungsidentifikation. Daher wird zur Reduzierung der Rechenzeit auf eine feinere Diskretisierung verzichtet. Der während der Identifikation betrachtete Zeitraum wird auf 23 Zeitschritte beschränkt. Diese Zeit benötigt die an der Scheibenoberseite hervorgerufene Kompressionswelle, um in einer ungeschädigten Struktur nach der Reflektion am gegenüberliegenden Rand wieder zu ihrem Ausgangspunkt zurückzukehren.

5.1.2 Anpassung des Kalman-Filters

Zur Schädigungsidentifikation wird das in Kapitel 3.2.2 beschriebene alternative Kalman-Filter benutzt. Die zu schätzenden Zustandsgrößen \mathbf{x}_k entsprechen den Parametern der zu lokalisierenden Schädigung aus dem Experiment. Es handelt sich hier um einen stationären, zeitinvarianten Prozeß, da sich die zu suchenden Parameter der Fehlstelle während der Optimierung nicht ändern. Die Transitionsmatrix Φ_k , die den Übergang des Systems aus dem Iterationsschritt t_k nach t_{k+1} beschreibt, entspricht daher der Einheitsmatrix \mathbf{I} . Daraus ergibt sich der Zusammenhang nach Gleichung (5.3).

Bei Durchführung von Labormessungen treten im allgemeinen Meßfehler auf, die das Kalman-Filter mit \mathbf{w}_k berücksichtigen kann. Da nur numerische Simulationen durchgeführt werden, werden diese Störgrößen zu Null gesetzt. Der Einfluß der Störung muß bei Messungen im Labor allerdings neu untersucht werden. Die Systemgleichung des Prozesses (3.19) vereinfacht sich zu

$$\mathbf{x}_{k+1} = \mathbf{x}_k \quad (5.3)$$

Der Meßwertvektor \mathbf{z}_k enthält die aus dem Experiment gewonnenen Verschiebungen auf dem äußeren Rand der Scheibe. Der Zusammenhang zwischen den aus der Randelementmethode berechneten Verschiebungen und den gesuchten Schädigungsparametern ist nichtlinear nach

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k, \quad (5.4)$$

wobei \mathbf{h} die Nichtlinearität beschreibt. Zur Schadenserkennung unter Verwendung des Kalman-Filters muß in jedem Iterationsschritt der Optimierung eine Linearisierung der Beziehung (5.4) stattfinden. Dies führt auf die Anwendung des erweiterten Kalman-Filters, in der die linearisierte Meßmatrix \mathbf{H}_k in jedem Iterationsschritt durch eine Sensitivitätsanalyse der Rißparameter bezogen auf die Verschiebungsgrößen neu bestimmt wird.

Der Vektor \mathbf{z}_k wird mit den aus der Simulation des Experimentes resultierenden Verschiebungen an jedem äußeren Diskretisierungspunkt zu jedem betrachteten Zeitschritt belegt. Für die oben gezeigte Diskretisierung der zweidimensionalen Scheibe folgt daraus ein Vektor mit 3.496 Komponenten. Im einfachsten Fall einer nur durch eine Schädigung gestörten Struktur enthält der Vektor \mathbf{x}_k nur vier Parameter. Desweiteren existiert zu Beginn der Optimierung keine Vorahnung über Lage und Form der Schädigung. Daher wird wie in Kapitel 3.2.3 beschrieben, die alternative Form des Kalman-Filters verwendet.

Somit läßt sich das Kalman-Filter wie folgt beschreiben:

0.) Voraussetzungen ($k = 0$):

- a) $\Phi_k = \mathbf{I}$
- b) $\hat{\mathbf{x}}_0^-$ und \mathbf{P}_0^- werden beliebig angenommen,

1.) Berechnung der Verformungen $\hat{\mathbf{z}}_k$ mit der Randelementmethode,

2.) Bestimmung von \mathbf{H}_k durch eine Sensitivitätsanalyse von $\hat{\mathbf{z}}_k$,

3.) Korrektur der Fehlerkovarianzmatrix

$$\mathbf{P}_k^{-1} = (\mathbf{P}_k^-)^{-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k, \quad (5.5)$$

4.) Ermittlung der Kalman-Verstärkungsmatrix zur Optimierung der Schätzung der Schädigungsparameter $\hat{\mathbf{x}}_k$,

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1}, \quad (5.6)$$

5.) Verbesserung der geschätzten Parameter $\hat{\mathbf{x}}_k$,

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_k + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k^-), \quad (5.7)$$

6.) Aktualisierung der Fehlerkovarianzmatrix

$$\mathbf{P}_{k+1}^- = \mathbf{P}_k + \mathbf{Q}_k, \quad (5.8)$$

7.) Erhöhe k um Eins und gehe zu 1.).

5.1.3 Identifikation einer Schädigung

Verformungen der Scheibe

Der Erfolg der Schadenslokalisierung hängt vom unterschiedlichen Verformungsverhalten einer geschädigten Struktur im Vergleich zu deren ungeschädigtem Original ab. Dieses abweichende mechanische Verhalten soll mit der Randelementmethode näher untersucht werden. Eine quadratische Scheibe mit mittig gelegenen Riß wird, wie in Abbildung 5.2 dargestellt, für die Rißerkennung betrachtet. In einer zweiten Simulation wird die gleiche Scheibe ohne Riß untersucht. Die Belastung bleibt in beiden Simulationen gleich und hat den zeitlichen Verlauf einer Heaviside-Funktion (Einheitssprungfunktion s. Kapitel 5.1.5 Gleichung (5.10)). In Abbildung 5.3 ist der zeitliche Verlauf der Verschiebung in Lastrichtung für den mittleren Punkt der Oberseite, gekennzeichnet als Punkt A in Abbildung 5.2, der geschädigten und der ungeschädigten Scheibe gegenübergestellt. Mit der Randelementmethode werden 100 Zeitschritte simuliert. Nach einer Zeit von ungefähr $t = 4\text{s}$ ist eine kleine Differenz der beiden Verformungen zu erkennen. Die durch den Riß hervorgerufene Wellenreflektion wird den betrachteten Meßpunkt auf der Scheibenoberseite zur Zeit

$$t = \frac{s}{c_1} = \frac{100}{23,44} = 4,26 \quad (5.9)$$

erreichen. Die von der Welle zurückgelegte Strecke wird in Gleichung (5.9) durch s bezeichnet. Die Kompressionswelle mit der Wellengeschwindigkeit c_1 wird wie erwartet an dem Riß reflektiert und führt zu der zur Schädigungsidentifikation notwendigen abweichenden Verformungsantwort der geschädigten Struktur. Die Randelementmethode liefert ein physikalisch plausibles Verformungsverhalten der Scheibe. Durch Minimierung der Verformungsdifferenz aller Meßpunkte am Außenrand der Scheibe scheint das Erkennen einer Schädigung innerhalb der Struktur möglich zu sein.

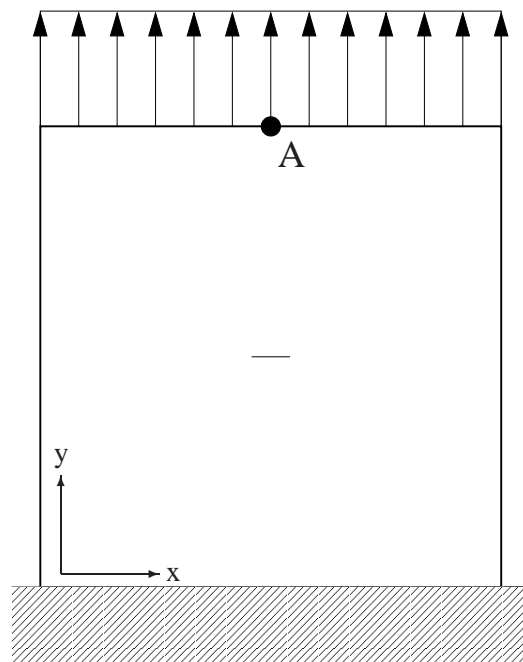


Abbildung 5.2: Belastung einer geschädigten Scheibe

Identifikation eines Risses

Beginnend mit einer einfachen Optimierungsaufgabe soll durch das Kalman-Filter ein rißförmiger Defekt in einer quadratischen Scheibe identifiziert werden. Die zu suchende Schädigung liegt nach Abbildung 5.2 in der Mitte der Scheibe, deren Oberseite mit einer Heaviside-Funktion belastet wird. Um einen anschaulichen Verlauf der Fehlerwertfunktion nach Gleichung (5.1) zu ermitteln, werden für mehrere willkürlich verteilte Rißpositionen die Fehlerwerte berechnet. Die Rißlänge und der Neigungswinkel des Risses werden dabei konstant vorgegeben und entsprechen den Werten des zu suchenden Risses. In Abbildung 5.4 ist der negative Fehlerwert in Abhängigkeit der Mittelpunktkoordinaten des Risses dargestellt. Es ergeben sich große Fehlerwerte, wenn ein Riß weiter weg vom zu suchenden Riß liegt. Somit werden für alle Rißpositionen am Rand der Scheibe hohe Fehlerwerte berechnet. Es ist anzunehmen, daß der Fehlerwert in Abhängigkeit der zu suchenden Koordinaten ein konvexes Verhalten aufweist. Deutlich ist der Anstieg des negativen Fehlerwertes zu seinem Maximum von $F = 0$ in der Mitte der Scheibe hin zu erkennen. Dieses konvexe Verhalten der Fehlerfunktion erleichtert die Lokalisierung der Schädigung. Dabei ist anzumerken, daß diese Konvexität nicht mathematisch bewiesen wird, und der Fehlerwert hier nur in Abhängigkeit der Mittelpunktkoordinaten eines Risses dargestellt wird.

Für die Identifikationsaufgabe benötigt das Kalman-Filter Startwerte. Bei Wahl von $\hat{\mathbf{x}}_0^- = [70 \ 20 \ 20 \ 1]$ als Startparameter so ergibt sich für die Suche nach dem in Abbildung 5.2 dargestellten Riß mit den Parametern $\mathbf{x} = [50 \ 50 \ 10 \ 0]$ der in Abbildung 5.5 gezeigte Optimie-

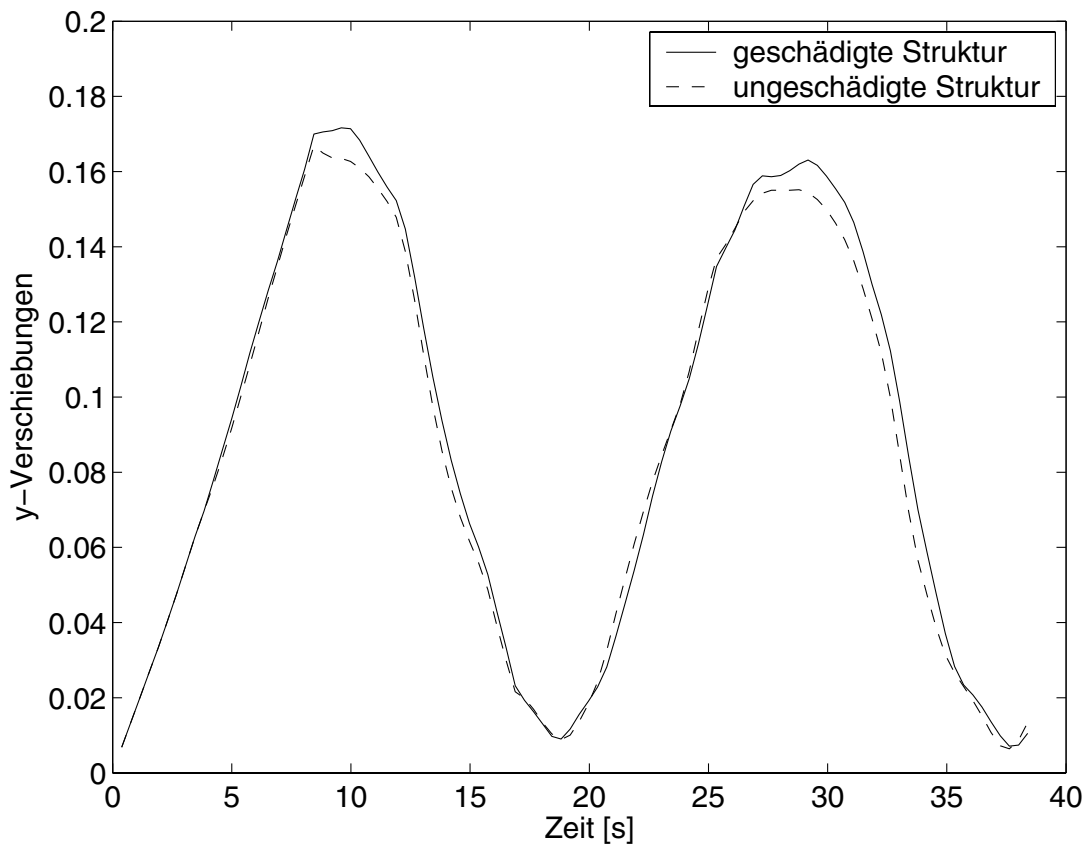


Abbildung 5.3: Verschiebung in Lastrichtung am Mittelpunkt der Scheibenoberseite

rungsverlauf. Die stark durchgezogenen Linien kennzeichnen die Zielwerte der Optimierung. Wie Abbildung 5.5 zeigt, wird die Position als auch die Länge und Ausrichtung des Risses nach 23 Iterationen sehr gut angenähert. Das Kalman-Filter stellt somit ein geeignetes Optimierungsverfahren zur Identifikation einer Schädigung dar. Der Fehlerwert unterschreitet am Ende der gezeigten Optimierung den Wert $1 \cdot 10^{-6}$. Andere Beispielrechnungen ergaben, daß bei dem Erreichen dieses Fehlerwertes die Parameter einer Schädigung genügend gut angenähert sind, so daß dieser im folgenden als Abbruchkriterium der Optimierung verwendet wird.

5.1.4 Einfluß der Schädigungsform

Zur Fehlstellenidentifikation werden zwei unterschiedliche Formen der Materialschädigung berücksichtigt. Zum einen eine ellipsenförmige Fehlstelle und zum anderen eine rißförmige Fehlstelle. In Abbildung 5.6 ist der Optimierungsverlauf für einen zu suchenden Riß mit den Parametern $\mathbf{x} = [50 \ 50 \ 10 \ 0]$ gezeigt. Die Startparameter sind $\hat{\mathbf{x}}_0^- = [30 \ 20 \ 8 \ 2]$. Nach 42 Iterationsschritten wird der Riß gefunden.

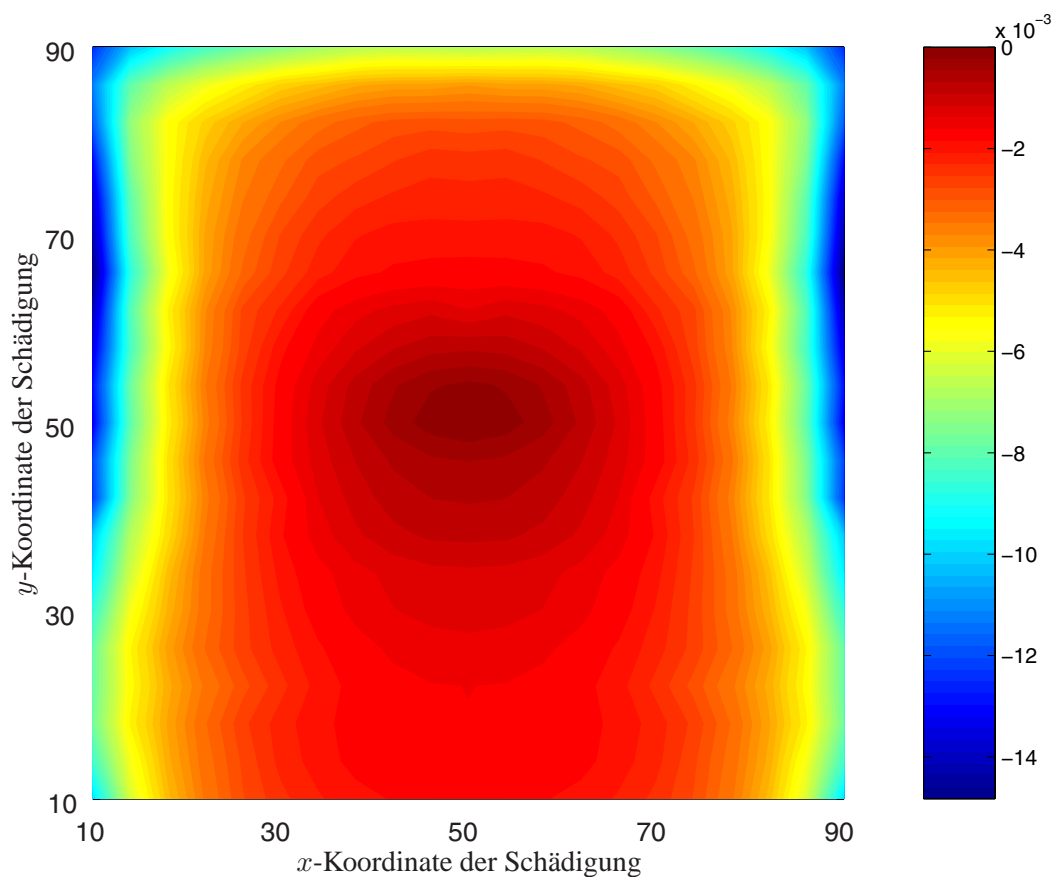
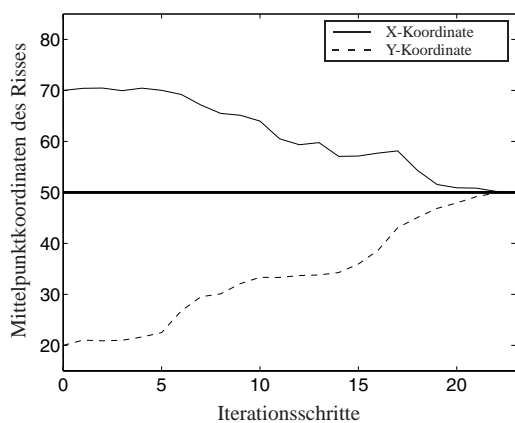
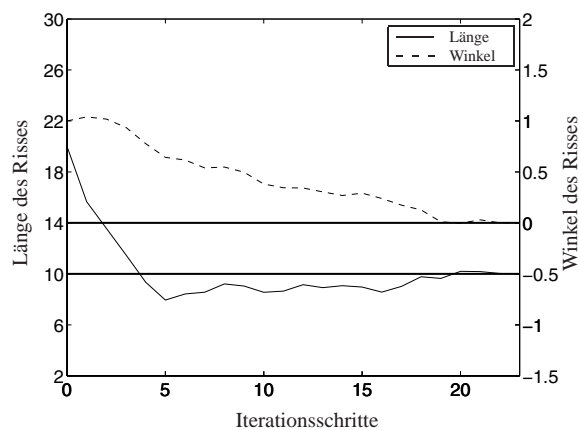


Abbildung 5.4: Fehlerwert in Abhängigkeit der Rißposition

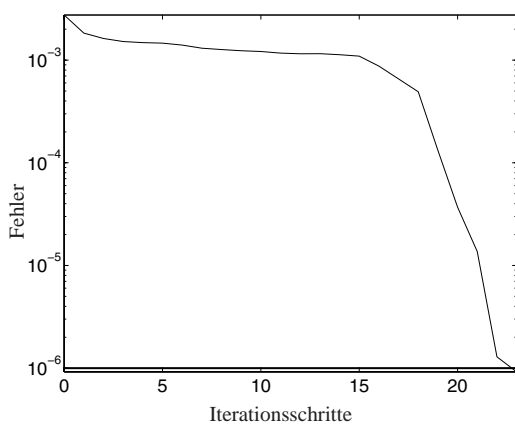
Wird dem gegenüber eine Ellipse mit ähnlichen Parametern $\mathbf{x} = [50 \ 50 \ 10 \ 10]$ und $\hat{\mathbf{x}}_0^- = [30 \ 20 \ 8 \ 12]$ gesucht, ergibt sich der in Abbildung 5.7 gezeigte Verlauf der Optimierung. Die ellipsenförmige Schädigung wird schon nach 8 Iterationsschritten identifiziert. Dieses Beispiel zeigt exemplarisch, daß Ellipsen einfacher zu lokalisieren sind als Risse, da sie auf Grund der von ihr verursachten Materialschwächung größere Deformationen der Scheibe hervorrufen. Daraus resultiert eine größere Abweichung der gemessenen und berechneten Verformungen der Scheibenaußenseite, die sich positiv auf die Konvergenzgeschwindigkeit der Optimierung durch das Kalman-Filter auswirken. Im folgenden wird nur noch die Identifikation von rißförmigen Schädigungen untersucht, wobei ein möglicher Kontakt der Rißoberflächen berücksichtigt wird.



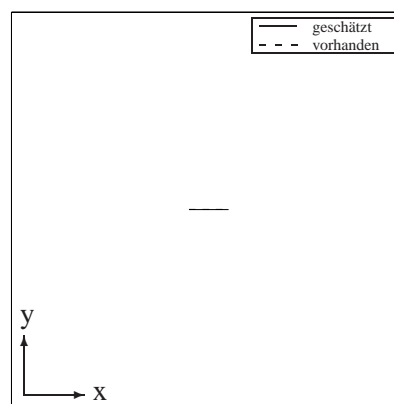
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

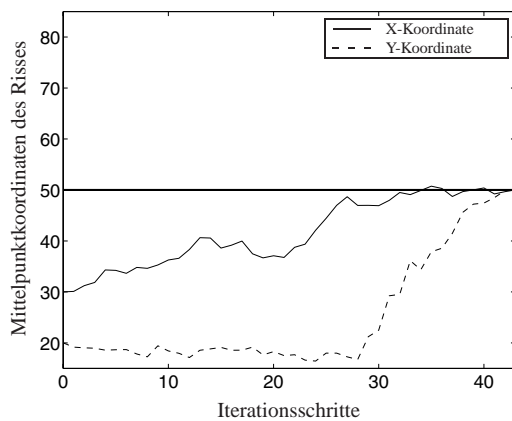


(c) Fehlerverlauf der Optimierung

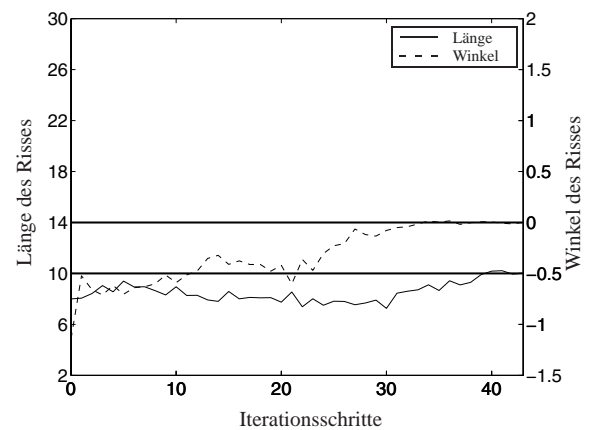


(d) Rißposition in der Scheibe

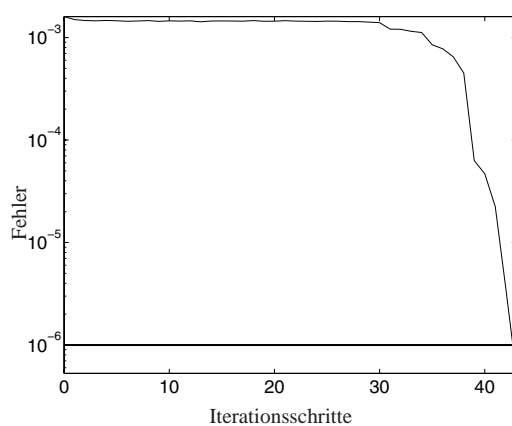
Abbildung 5.5: Iterationsverläufe und Geometrie für den Startwert [70 20 20 1]



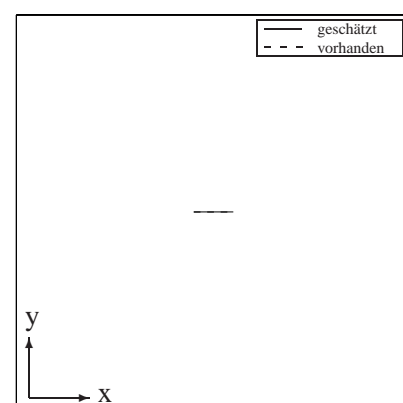
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

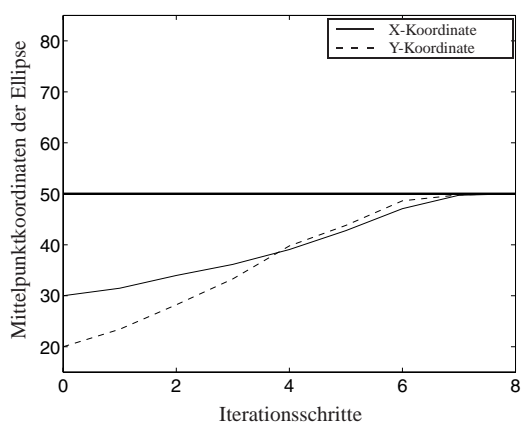


(c) Fehlerverlauf der Optimierung

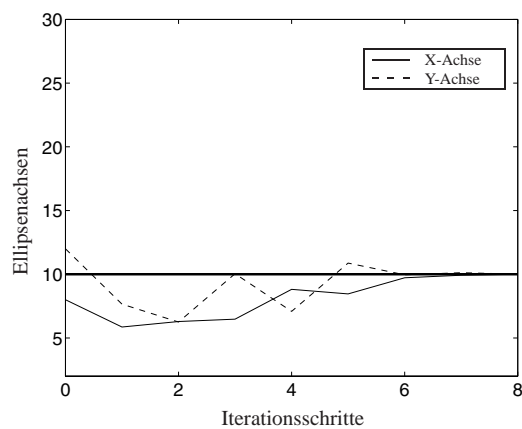


(d) Rißposition in der Scheibe

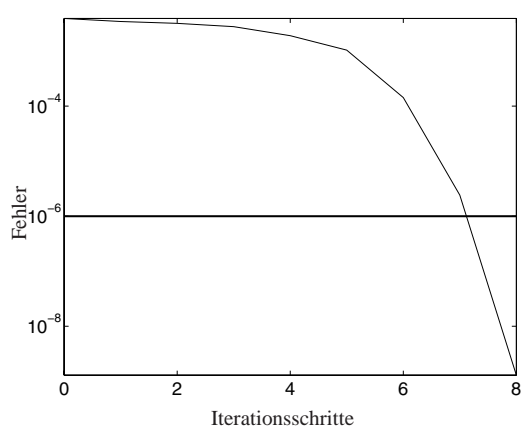
Abbildung 5.6: Identifikation eines Risses mit einem Startwert [30 20 8 2]



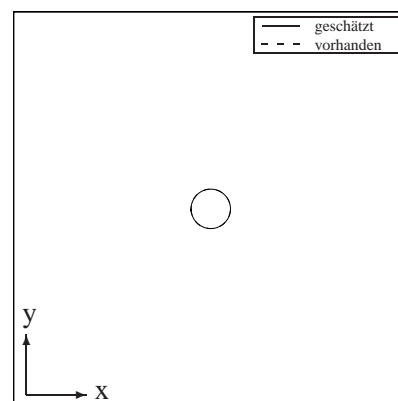
(a) Mittelpunktkoordinaten der Ellipse



(b) Ellipsenachsen



(c) Fehlerverlauf der Optimierung



(d) Ellipsenposition in der Scheibe

Abbildung 5.7: Identifikation einer Ellipse (hier der Sonderfall eines Kreises) mit einem Startwert [30 20 8 12]

5.1.5 Unterschiedliche Belastungsarten

Zeitlicher Verlauf der Belastung

Für die vorherigen Untersuchungen war die Belastung der Scheibe in Form einer Heaviside-Funktion aufgebracht worden. Es ist durchaus denkbar, daß andere Belastungsformen ebenfalls ein günstiges oder gar besseres Identifikationsergebnis hervorrufen können. Daher werden drei unterschiedliche Lasten zur Schadenserkennung für die in Abbildung 5.8a dargestellten Scheibe untersucht.

Heaviside-Funktion: Der zeitliche Verlauf der Heaviside-Funktion ist in Abbildung 5.8b skizziert. Über die gesamte Simulationszeit von 23 Zeitschritten wird eine konstante Last auf die Scheibenoberseite aufgebracht. Damit lautet die Funktion

$$H(t) = \begin{cases} 1 & \text{für } t > 0 \\ 0 & \text{für } t < 0 \end{cases} . \quad (5.10)$$

Ricker-Impuls: Der Ricker-Impuls stellt eine impulsartige Belastung dar, die im Gegensatz zum Dirac-Impuls keinen Sprung aufweist. In numerischen Simulationen kann bei Anwendung des Dirac-Impulses die Berechnung instabil werden. Zu dem ist der Ricker-Impuls innerhalb eines Meßlabors einfacher nachzustellen als der Dirac-Impuls. Der Ricker-Impuls ist definiert als

$$R(t) = a^2 e^{-\frac{a^2 t^2}{2}} (a^2 t^2 - 1) . \quad (5.11)$$

Der Koeffizient a wird für die Schadensidentifikation nach Testrechnungen zu $a = 1$ gewählt. Beginnend zu dem Zeitpunkt $t = 0$ ergeben sich aus Gleichung (5.11) zunächst negative Funktionswerte, die einer Druckbelastung entsprechen würden. Da aber nur Zugbelastungen untersucht werden sollen, muß der Funktionsverlauf so verschoben werden, daß er im Koordinatenursprung beginnt. Die Funktion ändert sich zu

$$R^*(t) = e^{-\frac{(t+1)^2}{2}} ((t+1)^2 - 1) \quad (5.12)$$

und ist in Abbildung 5.8c dargestellt.

Sinusfunktion: Als weitere Belastungsfunktion soll eine Sinus-Funktion angewandt werden. Ein Lastwechsel soll in sechs Zeitschritten durchlaufen werden. Somit ergibt sich folgende Lastfunktion

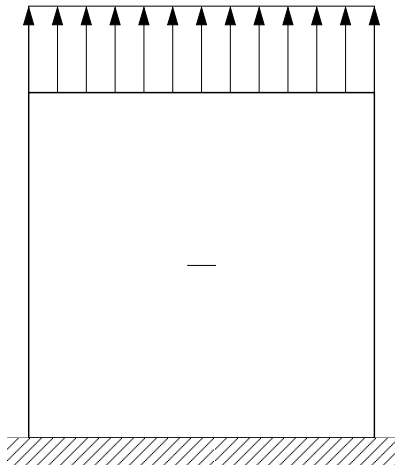
$$S(t) = \sin \left(\pi \frac{\Delta t + t}{6 \cdot \Delta t} \right) , \quad (5.13)$$

deren zeitlicher Verlauf in Abbildung 5.8d abgebildet ist.

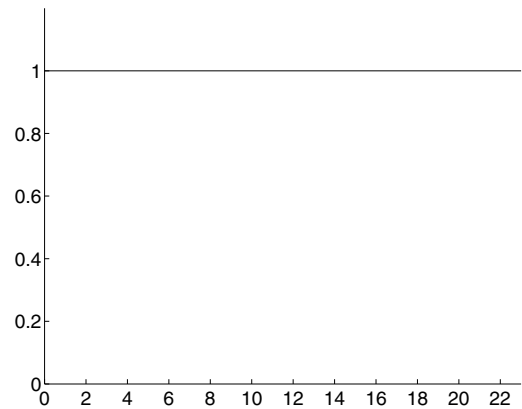
Alle Belastungsfunktionen werden mit einem Faktor 1.000 multipliziert. Die Verformungsantworten der Scheibe sind für den mittleren Punkt der Scheibenoberseite in Abbildung 5.9 in der linken Spalte dargestellt. Unterschieden wird dabei die Systemantwort der ungeschädigten

Struktur, die durch den gestrichelt skizzierten Verlauf abgebildet wird und der Verschiebungswerteverlauf der geschädigten Struktur, der mit einer durchgezogenen Linie wiedergegeben wird. Die Graphen verdeutlichen den physikalischen Effekt der Wellenreflektion. Nach Eintreffen der vom Riß reflektierenden Welle an der Oberfläche weichen die Verformungswerte der defekten Geometrie von der nicht fehlerhaften deutlich ab.

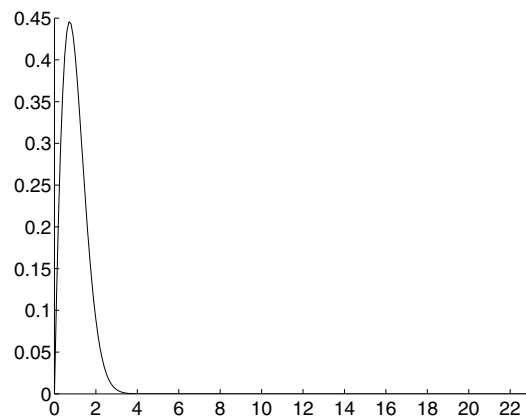
In der rechten Spalte von Abbildung 5.9 sind die negativen Fehlerwerte in Abhängigkeit der Rißposition abgebildet. Die Parameter der Länge und des Winkels werden konstant gehalten und haben die Werte der zu identifizierenden Schädigung. Während für die Heaviside- und Sinus-Belastung die Fehlerwerte in einem ähnlich großen Intervall liegen, sind die Fehlerwerte, berechnet unter Anwendung des Ricker-Impulses, auf einen engeren Wertebereich beschränkt. Nach Abbildung 5.9d hat der größtmögliche Betrag des Fehlerwertes einen Wert von nahezu null. Dieses Ergebnis deutet auf eine schlechtere Ausgangsposition für eine Schadenslokalisierung hin, falls die Scheibe mit einem Rickerimpuls belastet wird. Darüberhinaus ist ein zweites Extremum in der unteren Scheibengeometrie neben dem globalen Extremum der zu suchenden Rißposition erkennbar, wenn die Berechnung mit dem Ricker-Impuls oder der Sinusfunktion als Lastfunktion gewählt wird. Es besteht die Möglichkeit, daß sich die geschätzten Schädigungsparameter während der Optimierung dem lokalen Minimum annähern und von dort aus das globale Extremum der Fehlerfunktion nicht mehr erreichen können. Damit ist zu vermuten, daß die Identifikation mit der Belastung nach Gleichung (5.12) und (5.13) fehlschlagen wird.



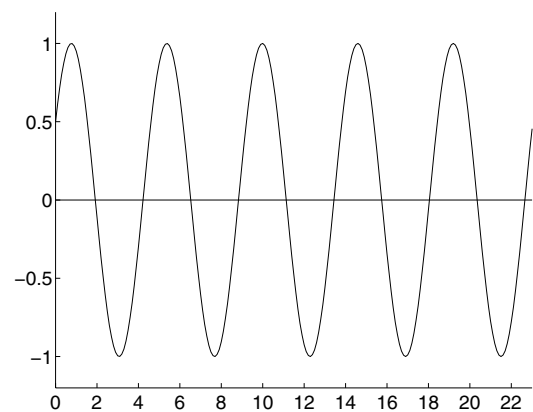
(a) Dynamische Belastung einer rechteckigen Scheibe



(b) Heaviside-Funktion

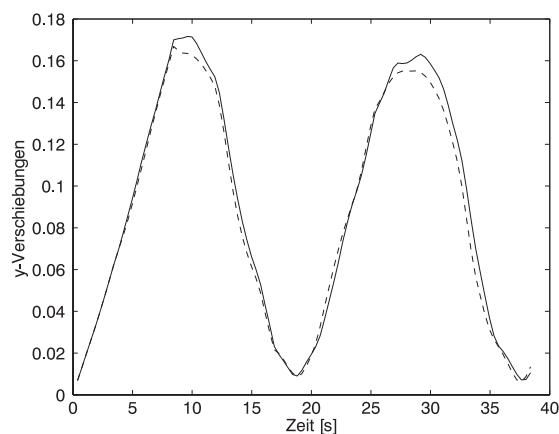


(c) Ricker-Impuls

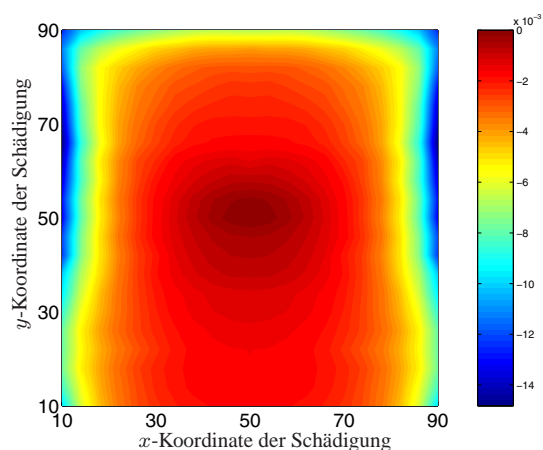


(d) Sinus-Funktion

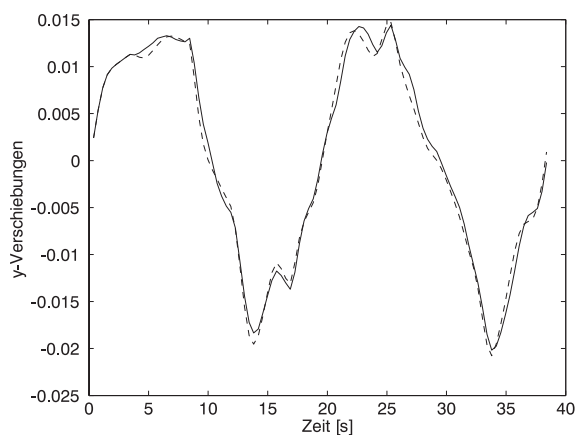
Abbildung 5.8: Getestete Belastungsfunktionen zur Identifizierung eines Risses mit dem Kalman-Filter



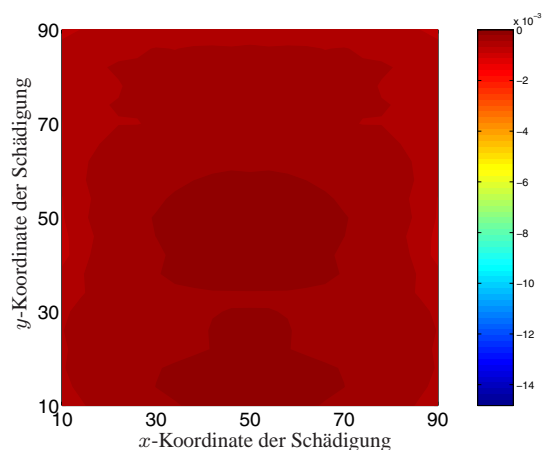
(a) Verschiebungen infolge Heaviside-Funktion



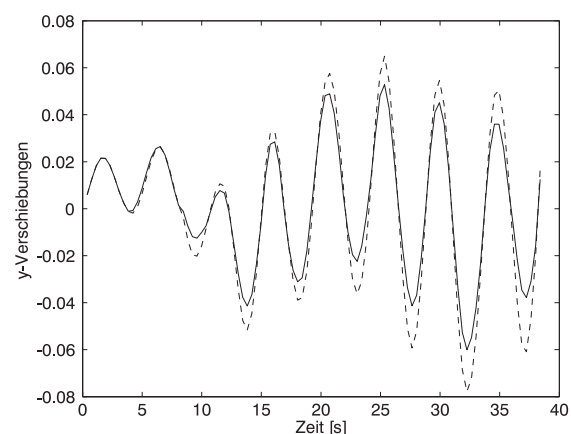
(b) Fehlerwerte bei unterschiedlichen Rißpositionen bei Heaviside-Belastung



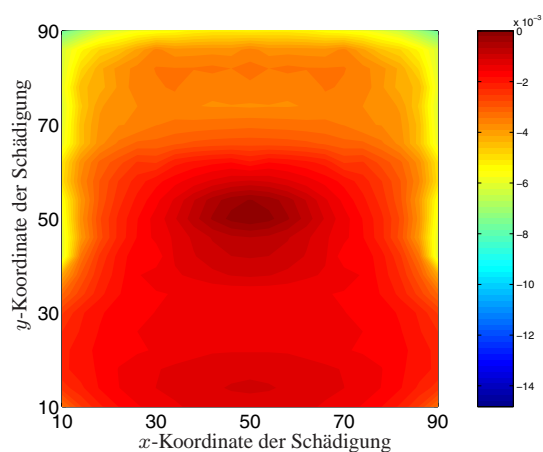
(c) Verschiebungen infolge Ricker-Impuls



(d) Fehlerwerte bei unterschiedlichen Rißpositionen bei Ricker-Belastung

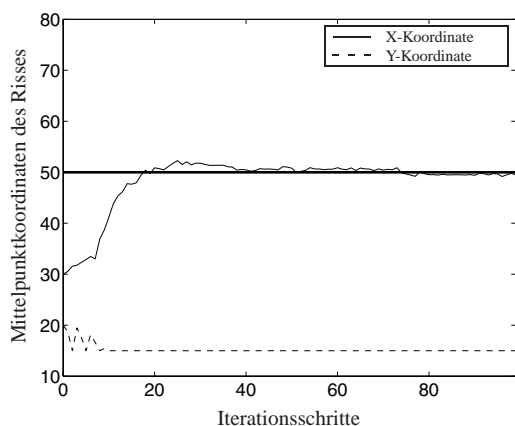


(e) Verschiebungen infolge Sinus-Funktion

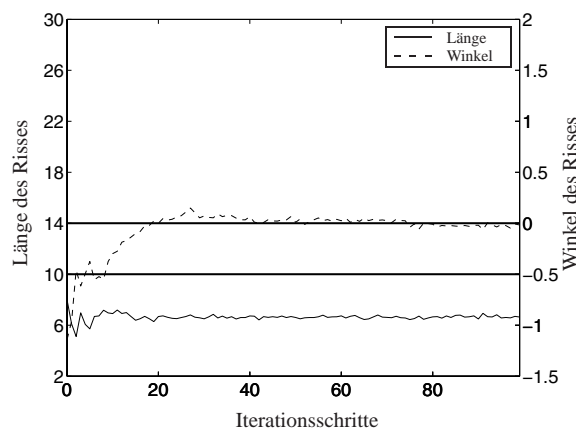


(f) Fehlerwerte bei unterschiedlichen Rißpositionen bei Sinus-Belastung

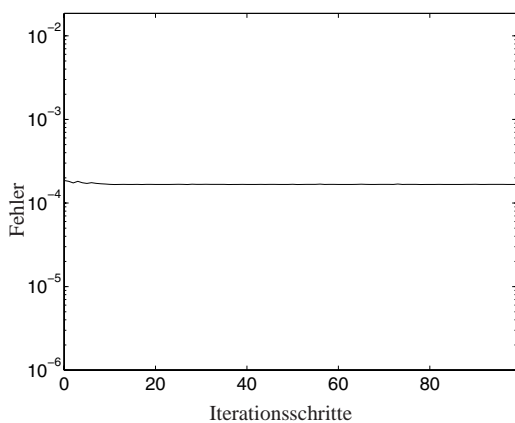
Abbildung 5.9: Verschiebungen der Scheibenoberseite und Fehlerwerte in Abhängigkeit der Rißkoordinaten bei unterschiedlichen Belastungsfunktionen, Vorgabe der Rißparameter durch [50/50/10/0]



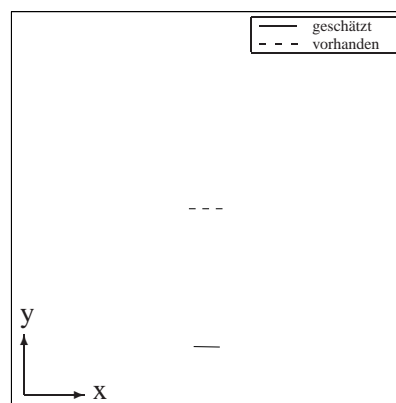
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



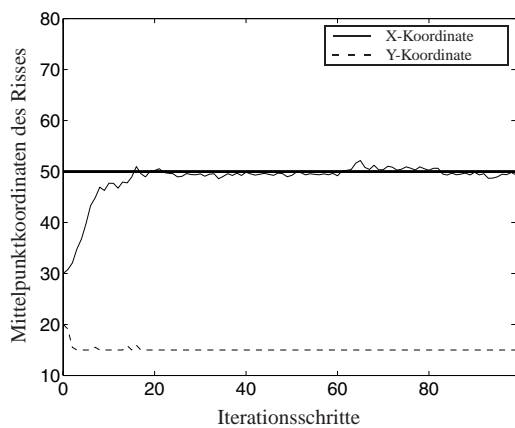
(c) Fehlerverlauf der Optimierung



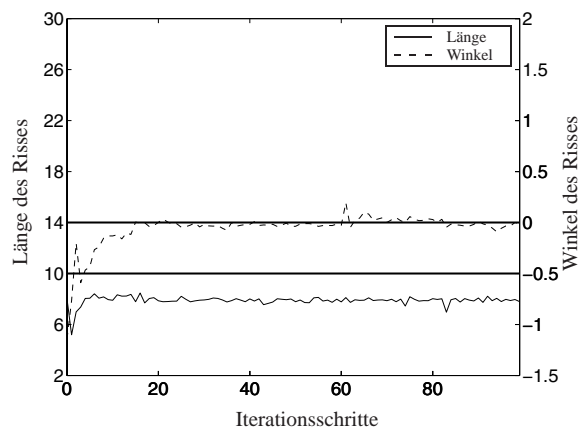
(d) Rißposition in der Scheibe

Abbildung 5.10: Identifikation eines Risses durch eine Ricker-Impuls Belastung mit dem Kalman-Filter und einem Startwert $[30 \ 20 \ 8 \ 2]$

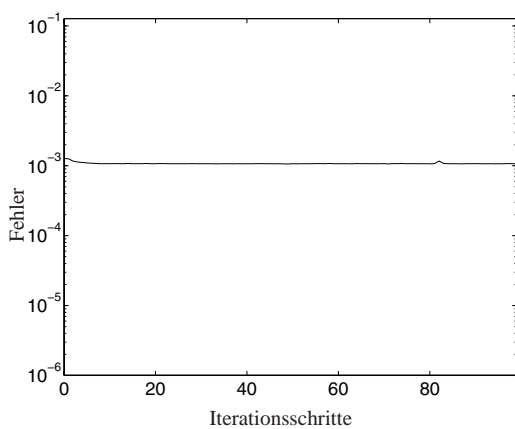
In den Abbildungen 5.10 und 5.11 ist der Verlauf der Optimierung bei Anwendung des Ricker-Impulses (5.12) und der Sinus-Funktion (5.13) abgebildet. Die Startwerte sind zu $\hat{\mathbf{x}}_0^- = [30 \ 20 \ 8 \ 2]$ gewählt, während die zu suchenden Schädigungsparameter mit $\mathbf{x} = [50 \ 50 \ 10 \ 0]$ vorgegeben werden. Der Identifikationsvorgang schlägt fehl, da sich der Optimierungsalgorithmus im lokalen Minimum in der unteren Scheibenhälfte festsetzt. Deutlich ist die Übereinstimmung der geschätzten Rißposition mit den Extrema in Abbildung 5.9 zu erkennen. Das Abbruchkriterium des Fehlerwertes wird in keinem der beiden Fälle erreicht. Die Rißidentifikation verläuft allerdings erfolgreich unter Gebrauch der Heaviside-Funktion wie in Abbildung 5.6 gezeigt wurde. Ausgehend von den gleichen Startwerten werden alle Rißparameter eindeutig identifiziert.



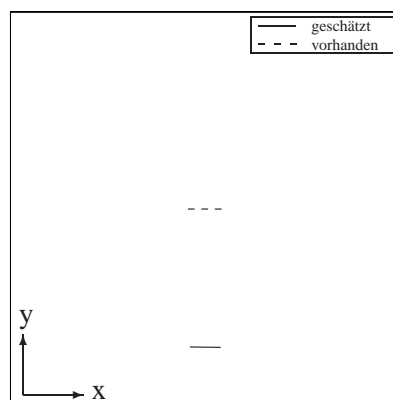
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



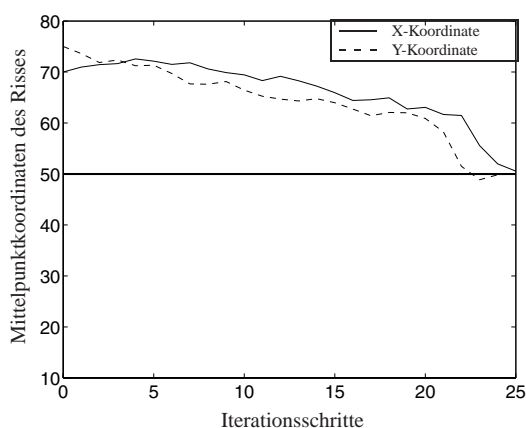
(c) Fehlerverlauf der Optimierung



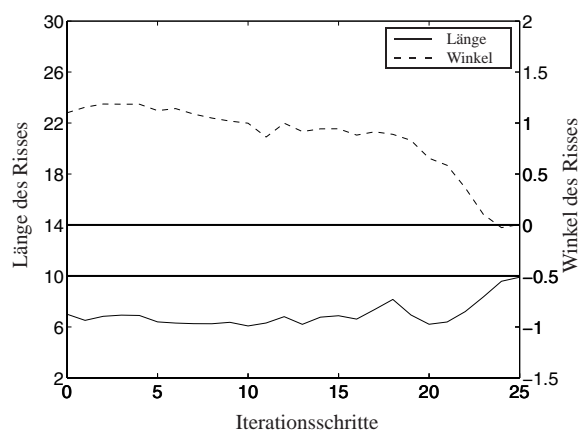
(d) Rißposition in der Scheibe

Abbildung 5.11: Identifikation eines Risses durch eine Sinus Belastung mit dem Kalman-Filter und einem Startwert [30 20 8 2]

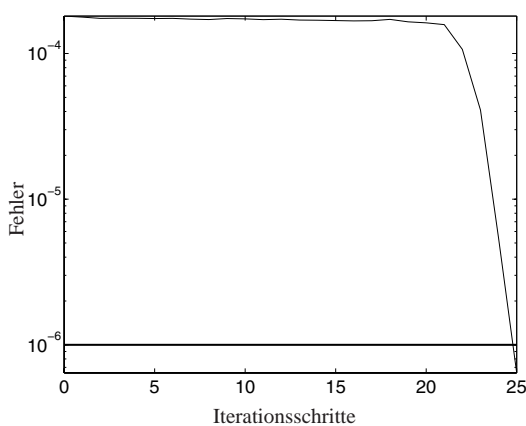
Abbildungen 5.12, 5.13 und 5.14 zeigen die Optimierungsverläufe unter Verwendung der drei untersuchten Belastungsfunktionen für die Startwerte $\hat{\mathbf{x}}_0^- = [70 \ 75 \ 7 \ 1, \ 1]$. In allen drei Fällen wird die Schädigung erfolgreich aufgedeckt. Für die Ricker-Impuls-Belastung wird das Abbruchkriterium erst nach 25 Iterationen erreicht, während der Optimierungsvorgang mit der Sinus- und Heaviside-Funktion nach 15 Iterationsschritten erfolgreich beendet wird. Dieses Ergebnis deutet darauf hin, daß der vorhin erwähnte eingeschränkte Wertebereich des Fehlerwertes nach Abbildung 5.9(d) die Konvergenzgeschwindigkeit des Optimierungsalgorithmus herabsetzt und die benötigte Zeit für den Identifikationsvorgang damit verlängert.



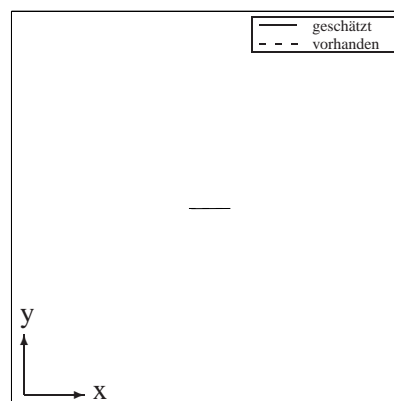
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

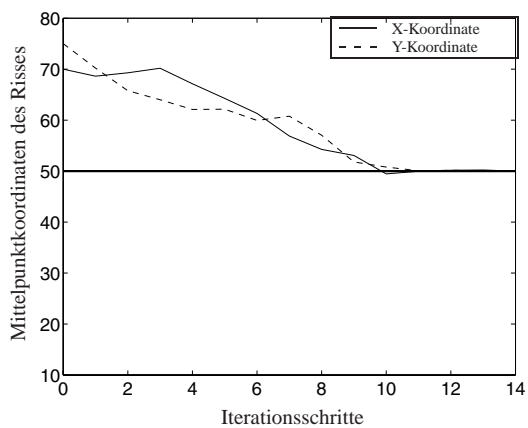


(c) Fehlerverlauf der Optimierung

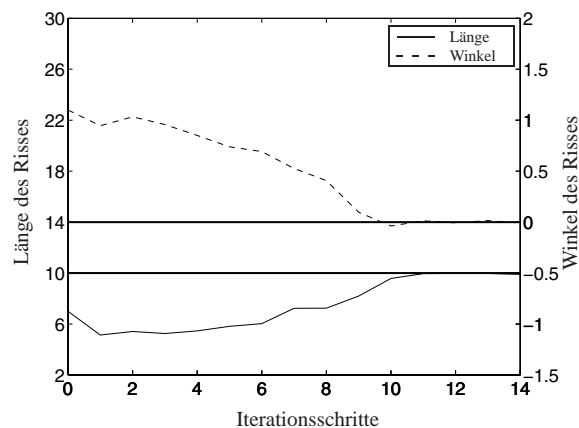


(d) Rißposition in der Scheibe

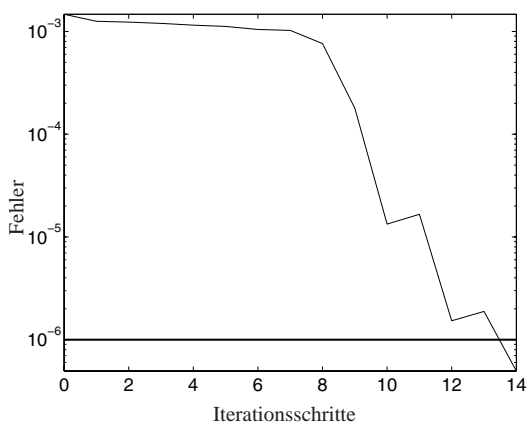
Abbildung 5.12: Identifikation eines Risses durch eine Ricker-Impuls Belastung mit dem Kalman-Filter und einem Startwert $[70 \ 75 \ 7 \ 1, \ 1]$



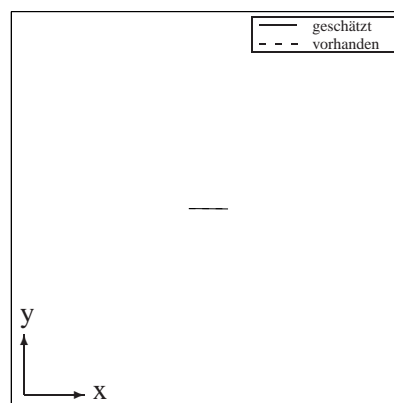
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



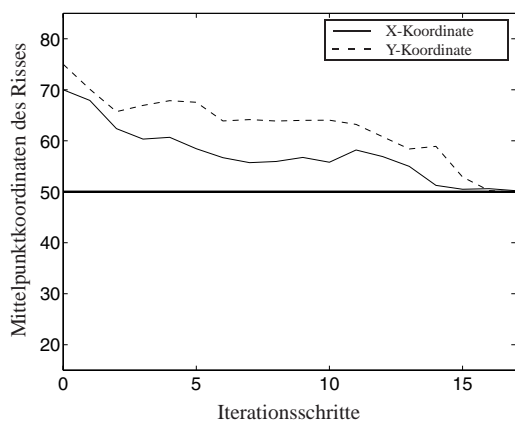
(c) Fehlerverlauf der Optimierung



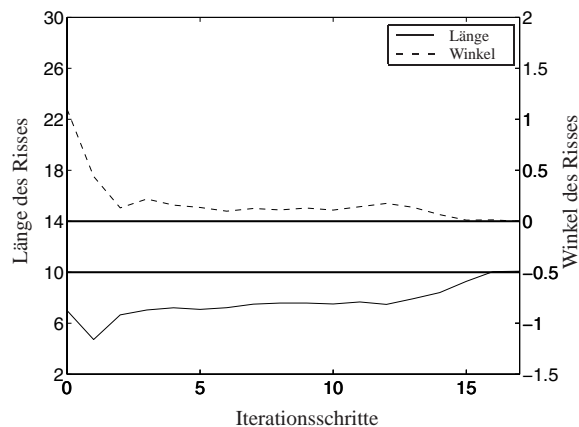
(d) Rißposition in der Scheibe

Abbildung 5.13: Identifikation eines Risses durch eine Sinus Belastung mit dem Kalman-Filter und einem Startwert [70 75 7 1,1]

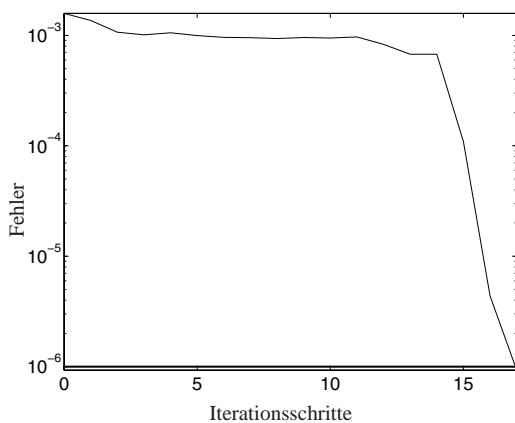
Die Ergebnisse der Untersuchungen deuten darauf hin, daß zur erfolgreichen Fehlstelleniden-



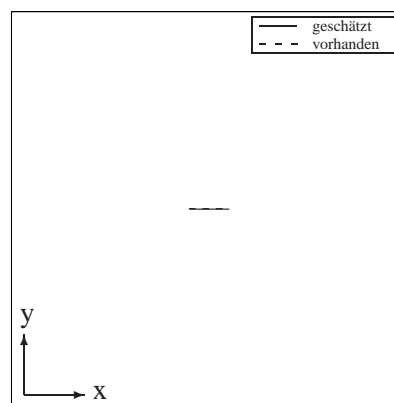
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



(c) Fehlerverlauf der Optimierung



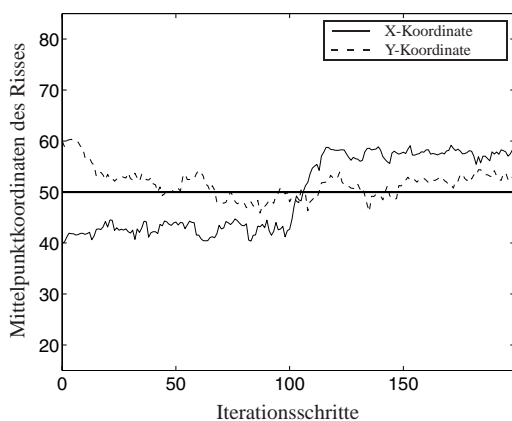
(d) Rißposition in der Scheibe

Abbildung 5.14: Identifikation eines Risses durch eine Heaviside-Belastung mit dem Kalman-Filter und einem Startwert [70 75 7 1,1]

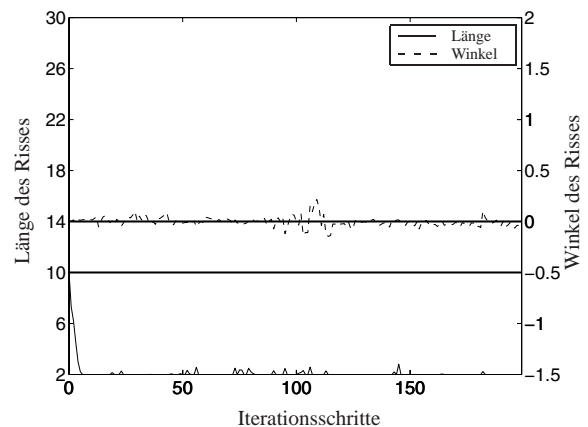
tifikation die Scheibe mit einer Heaviside-Funktion belastet werden sollte. Die zwei anderen hier vorgestellten Belastungsfunktionen sind für geeignete Startwerte auch zur Identifikation von Schädigungen verwendbar, weisen jedoch eine langsamere Konvergenz auf und haben zusätzlich ein lokales Extremum innerhalb des Fehlerwertverlaufes.

Zug- oder Druckbelastung

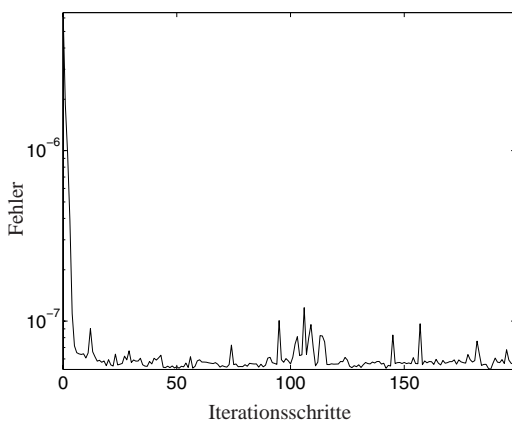
Für die bisherigen Untersuchungen ist eine dynamische Zugbelastung aufgebracht worden. Unter der Annahme, daß eine dynamische Belastung in Form einer Heaviside-Funktion am geeignetsten für die Schadenserkennung ist, wird in diesem Abschnitt das Identifikationsverhalten durch eine Druckbelastung näher betrachtet. Es sei hier nochmal darauf hingewiesen, daß bei Belastung der Geometrie durch eine Druckbelastung der Kontaktalgorithmus bei Berechnung der resultierenden Verformungen aktiviert wird. Die schon zuvor untersuchte, geschädigte Scheibe wird mit einer Heaviside-Druckbelastung angeregt. Abbildung 5.15 zu Folge, können



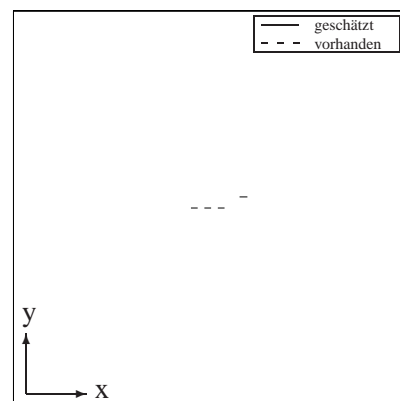
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



(c) Fehlerverlauf der Optimierung

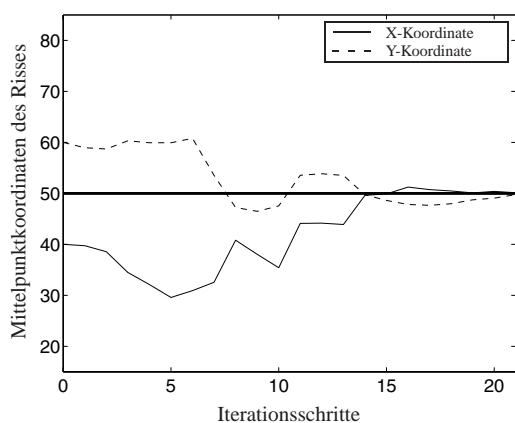


(d) Rißposition in der Scheibe

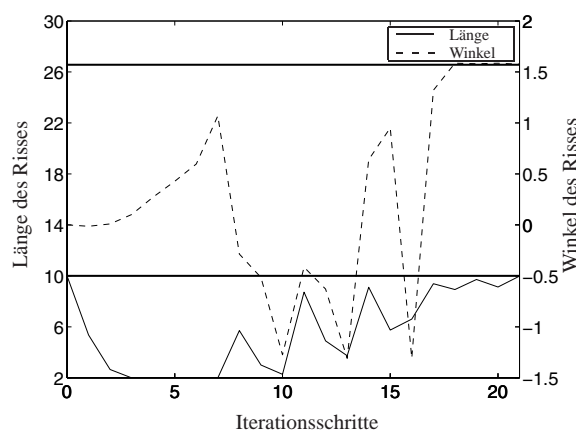
Abbildung 5.15: Druckbelastung: Iterationsverläufe und Geometrie für den Startwert [40 60 10 0]

die Rißparameter nicht erkannt werden. Auffallend ist hier die Identifikation eines sehr kurzen Risses. Dies deutet im Prinzip auf eine ungeschädigte Geometrie hin. Der Einfluß des Risses auf das Systemverhalten der Scheibe unter Druckbelastung ist sehr gering. Dieser Sachverhalt

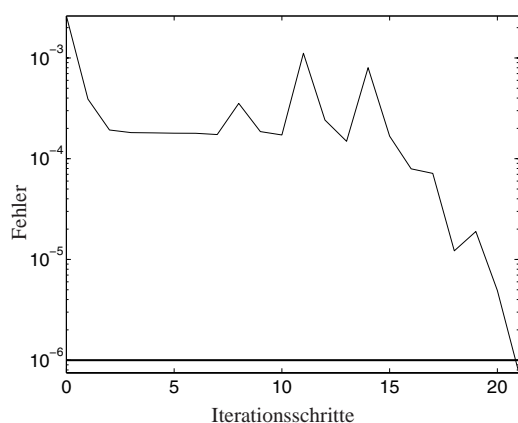
wird durch die niedrigen Fehlerwerte in Abbildung 5.15c bestätigt.



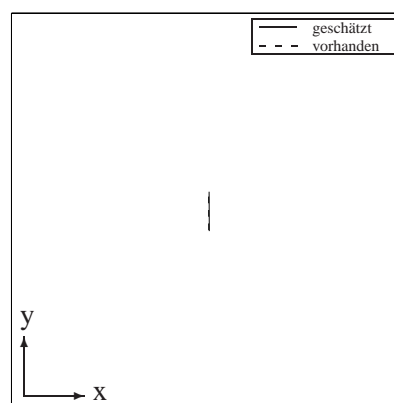
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



(c) Fehlerverlauf der Optimierung



(d) Rißposition in der Scheibe

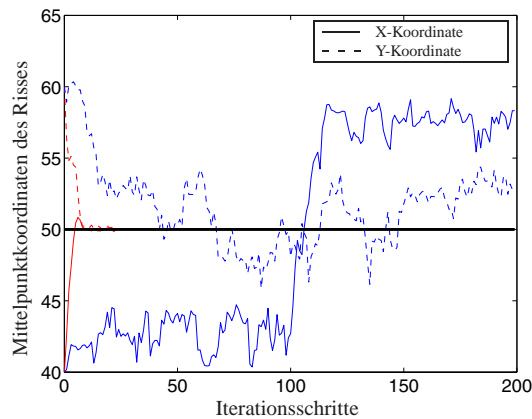
Abbildung 5.16: Zugbelastung: Iterationsverläufe und Geometrie für den Startwert [40 60 10 0]

Der in Abbildung 5.15 gezeigte Optimierungsverlauf deutet darauf hin, daß eine Druckbelastung in einigen Fällen eine Schadensidentifikation ungünstig beeinflussen kann, da die Fehlstelle kaum Auswirkungen auf die Verschiebungen des äußeren Scheibenrandes hat.

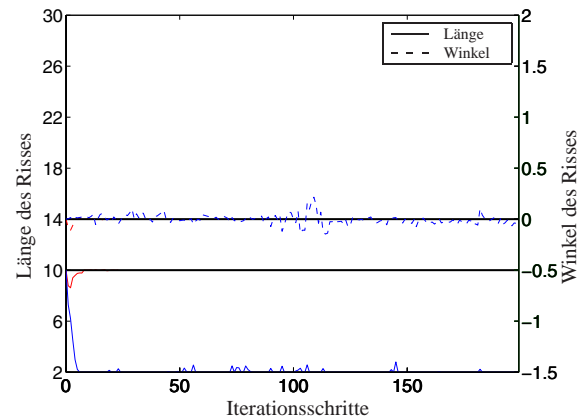
In Abbildung 5.16d wird ein weiteres Beispiel der Rißerkennung gezeigt. Das Kalman-Filter soll ein zu der eingespannten Scheibenseite senkrecht liegenden Riß durch eine Zugbelastung identifizieren. Im Allgemeinen ist hier davon auszugehen, daß der Riß sich auch in dieser Simulation nicht oder nur sehr wenig öffnet und damit der Identifikationsvorgang erschwert werden könnte. Abbildung 5.16 zeigt jedoch, daß der Riß mit einer dynamischen Zugbelastung schon nach 21 Iterationen identifiziert werden kann.

Der Optimierungsverlauf der Schädigungssituation aus Abbildung 5.15d ist in Abbildung 5.17 einmal mit Verwendung einer dynamischen Zugbelastung und zum anderen durch Anwendung

einer Druckbelastung gegenübergestellt. Es werden für beide Optimierungsvorgänge die gleichen Startwerte gewählt. Durch die Zugbelastung werden die Rißparameter schon nach wenigen Iterationen gefunden. Nach 200 Iterationen sind die Schätzungen mit Druckbelastung immer noch ungenügend.



(a) Mittelpunktkoordinaten des Risses

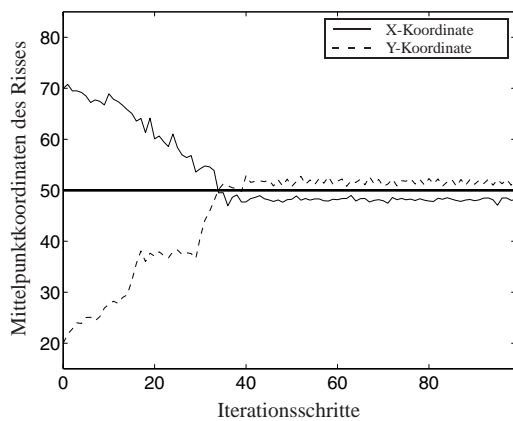


(b) Länge und Winkel des Risses

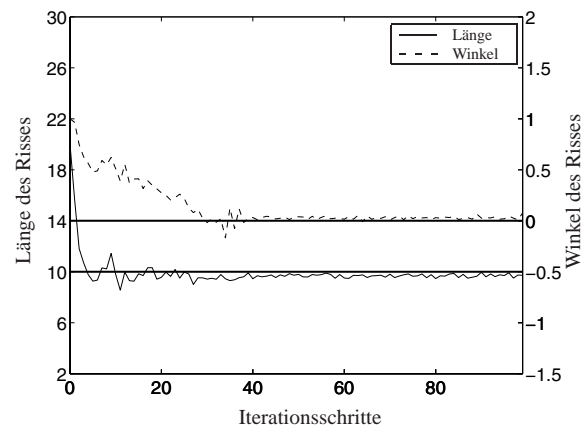
Abbildung 5.17: Iterationsverläufe bei Identifikation eines Risses im Vergleich: — Zug , — Druck

Es kann abschließend festgestellt werden, daß die Identifikation einer rißförmigen Schädigung in einer auf Druck belasteten Scheibe nicht für alle Ausrichtungen des Risses möglich ist. Die Optimierung konvergiert hingegen bei einer Zugbelastung auch bei der in Abbildung 5.15d gezeigten Schädigungssituation. Eine Belastung auf Zug ist zur Schädigungsidentifikation einer Druckbelastung vorzuziehen. Es werden daher im folgenden nur dynamische Zugbelastungen untersucht.

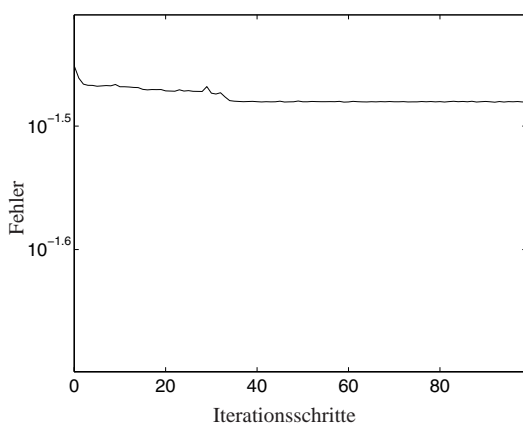
5.1.6 Schädigungsidentifikation mit verrauschten Meßgrößen



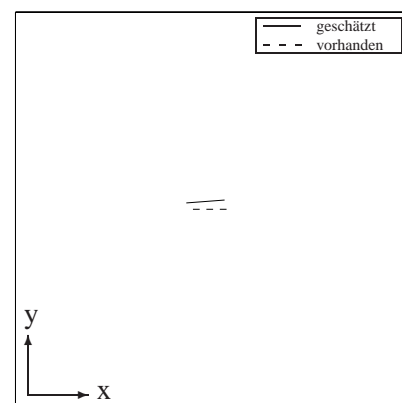
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



(c) Fehlerverlauf der Optimierung



(d) Rißposition in der Scheibe

Abbildung 5.18: Identifikation eines Risses mit dem Kalman-Filter mit bis zu 11 % fehlerbehafteten Eingangsgrößen

Die im Labor gemessenen Verformungen eines geschädigten Körpers werden unter anderem auf Grund von Meßfehlern nie exakt mit den Ergebnissen der numerischen Simulation übereinstimmen. Nur bei numerischen Simulationen der Verformungen sind die berechneten Verformungswerte bei richtig identifizierten Rißparametern identisch mit denen des „Experimentes“. Es wird deshalb das Verhalten der Optimierung unter der Bedingung untersucht, daß die gemessenen Verformungen des „Experimentes“ mit einem zufälligen Fehler behaftet sind. Es werden die exakten Daten berechnet, und diese werden anschließend mit einer zufällig bestimmten Größe leicht verändert. Dabei darf die zufällig bestimmte Größe einen bestimmten prozentualen Wert der exakten Verformungsdaten nicht überschreiten. In Abbildung 5.18 ist die Optimierung für einen zufälligen Fehler der Verformungen im Intervall von ± 11 Prozent für 100 Iterations-

schritte dargestellt. Der zu suchende Riß wird zwar nicht exakt gefunden, ist aber für das relativ große Fehlerintervall gut angenähert. In Abbildung 5.19a sind die Optimierungsverläufe für

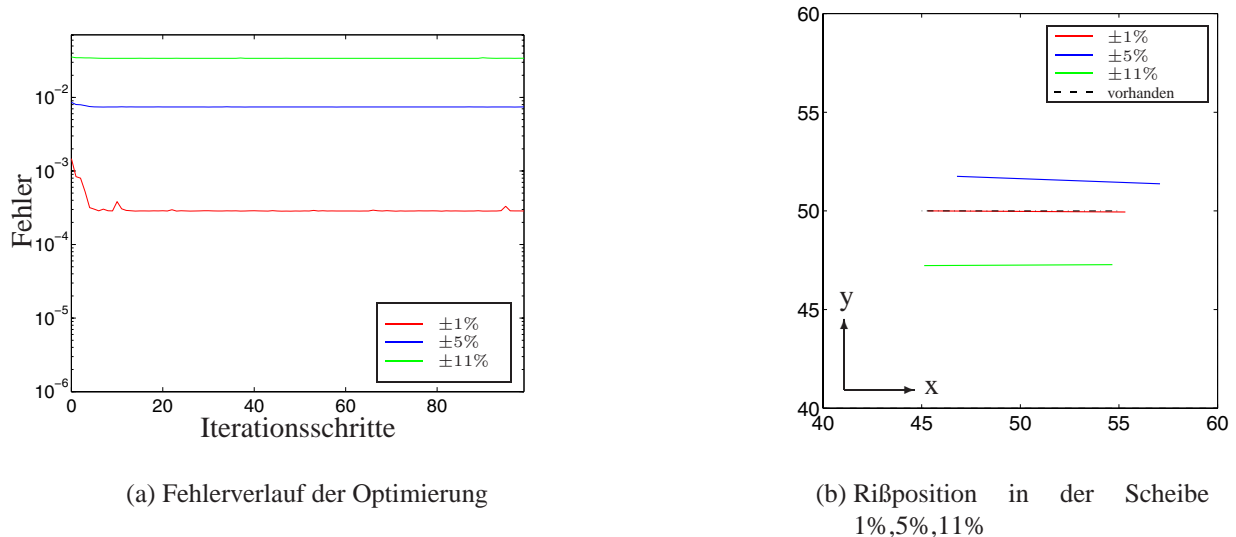


Abbildung 5.19: Fehlerverläufe und Ausschnitt der Geometrie für verschiedene Meßfehlerintervalle bei einer Identifizierung eines Risses durch das Kalman-Filter

weitere fehlerbehaftete Meßgrößen dargestellt. Die Verformungswerte des Experiments werden mit einem maximalen Fehler von 1%, 5% und mit 11% beaufschlagt. Ein Abbruchkriterium für den Fehlerwert wird zwar nicht erreicht, doch wird der Riß nach Abbildung 5.19b gut erkannt. Für einen Fehler der „gemessenen“ Verformungswerte im Bereich ± 1 Prozent werden die Schädigungsparameter sogar exakt bestimmt.

Die Untersuchungen haben gezeigt, daß der Kalman-Filter trotz leicht fehlerhaften Meßgrößen eine gute Identifizierung der Schädigung erzielen kann.

5.1.7 Identifikation mehrerer Schädigungen

Bisher wurden nur Scheiben mit einer Schädigung betrachtet. In diesem Abschnitt wird das Identifikationsverhalten von mehrfach geschädigten Strukturen untersucht. Ein weiterer Lastfall soll den Identifikationsvorgang beschleunigen.

5.1.8 Einfluß eines zweiten Lastfalls

Bei der Betrachtung von mehr als einer Schädigung in einer Scheibe kann es vorkommen, daß wie schon in Kapitel 4.4.1 in der Statik untersucht, die Auswirkung einer Schädigung auf die Verformung des Randes durch eine andere Schädigung überlagert wird. Eine derartige Schädigungskonstellation tritt auf, wenn zwei Fehlstellen in Lastrichtung gesehen übereinander liegen. Die Vermutung soll bestätigt werden, daß durch die Hinzunahme eines zweiten Lastfalls die Überlagerung der Einflüsse auf den Rand kompensiert werden kann. Zur Untersuchung dieser Vermutung wird eine Schädigungssituation mit zwei untereinander liegenden Rissen zunächst mit dem in Abbildung 5.20a skizzierten Lastfall simuliert. Zum Vergleich wird die gleiche geschädigte Struktur mit beiden Lastfällen nach Abbildung 5.20 belastet. Die Hinzunahme des

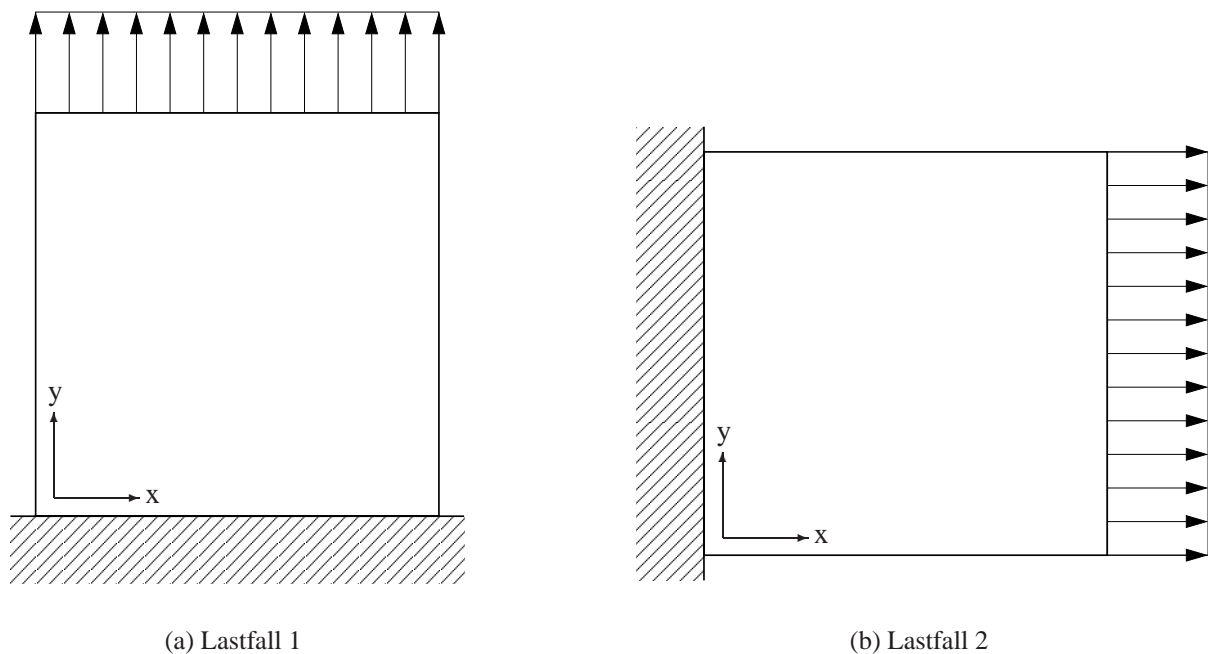
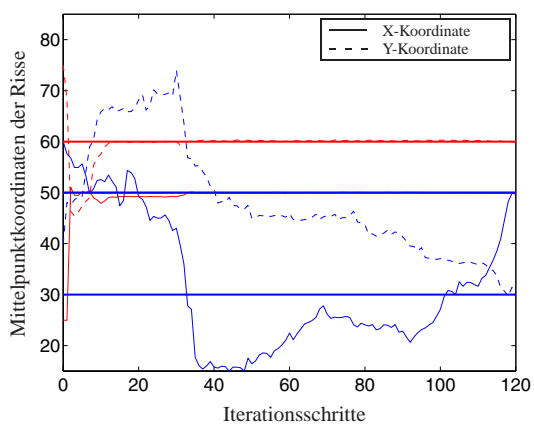
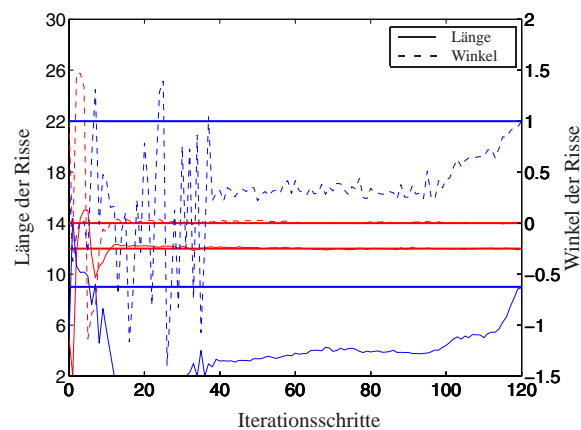


Abbildung 5.20: Die betrachteten Belastungsfälle

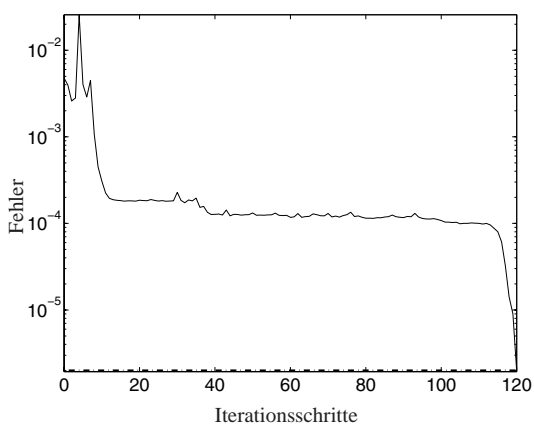
zweiten Lastfalls bewirkt eine Verdopplung der Anzahl der Komponenten im Meßvektor $\hat{\mathbf{z}}_k$. Erst nach Berechnung beider Lastfälle wird die nächste Schätzung durch das Kalman-Filter vorgenommen. Durch die Hinzunahme des zweiten Lastfalls verdoppelt sich daher die Rechenzeit pro Iterationsschritt.



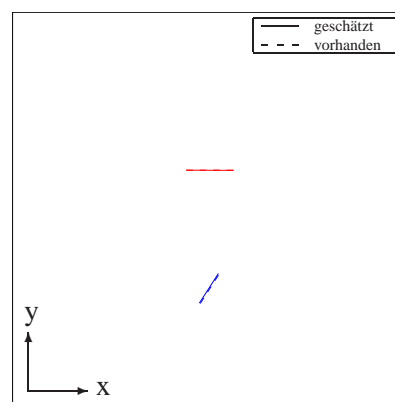
(a) Mittelpunktkoordinaten der Risse



(b) Länge und Winkel der Risse



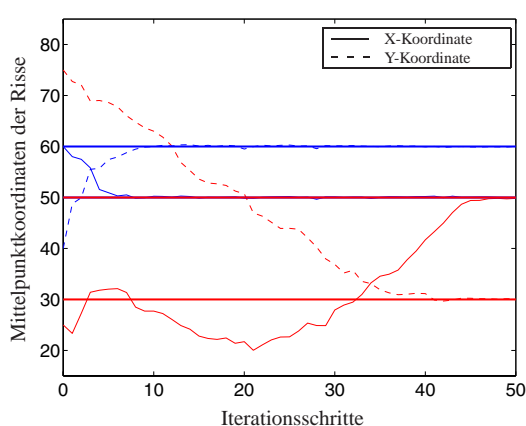
(c) Fehlerverlauf der Optimierung



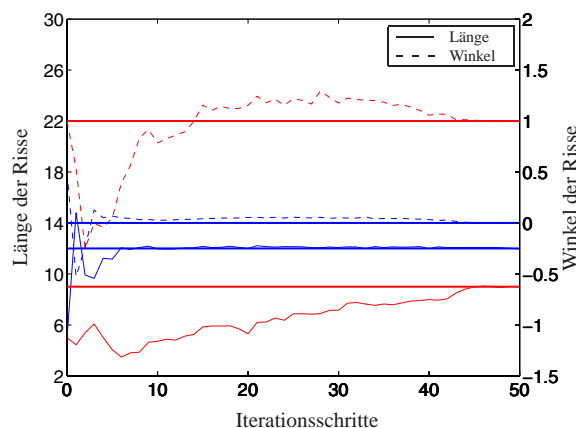
(d) Rißposition in der Scheibe

Abbildung 5.21: Identifikation von zwei Rissen mit dem Kalman-Filter durch einen Lastfall

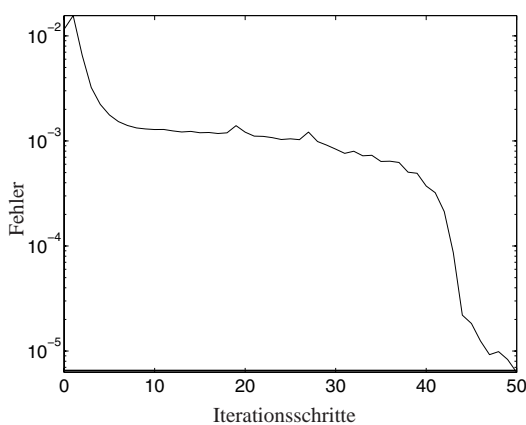
In den Abbildungen 5.21 und 5.22 sind die Optimierungsverläufe für die gleiche Schädigungssituation bei Verwendung eines Lastfalls und zweier Lastfälle dargestellt. Die Optimierung mit nur einem Lastfall benötigt 120 Iterationsschritte, während die mit zwei Lastfällen hingegen nach 50 Schritten abgeschlossen ist. In beiden Rechnungen werden die Risse identifiziert, die Konvergenzgeschwindigkeit erhöht sich durch die Hinzunahme des zweiten Lastfalls. In dem betrachteten Fall rentiert sich die Erhöhung der Rechenzeit pro Iterationsschritt durch den zweiten Lastfall. Rechnungen bei anderen Schädigungskonstellationen ergaben jedoch nicht immer eine Reduzierung der Gesamtrechenzeit.



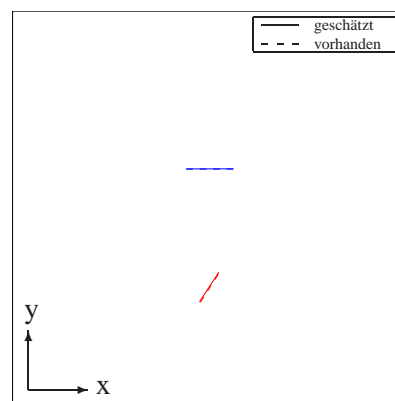
(a) Mittelpunktkoordinaten der Risse



(b) Länge und Winkel der Risse



(c) Fehlerverlauf der Optimierung

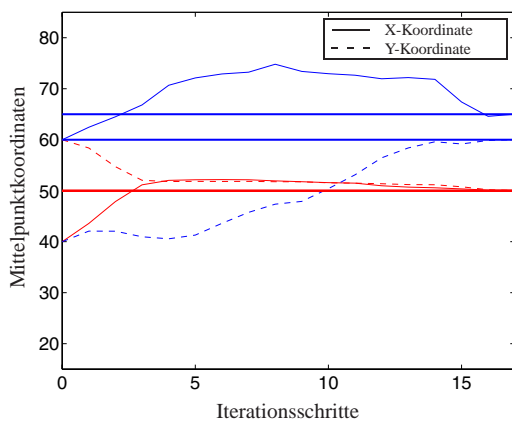


(d) Rißposition in der Scheibe

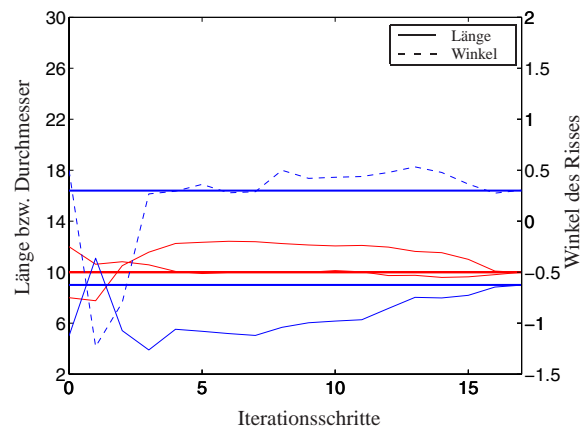
Abbildung 5.22: Identifikation von zwei Rissen mit dem Kalman-Filter durch zwei Lastfälle

5.1.9 Einfluß unterschiedlicher Schädigungsformen

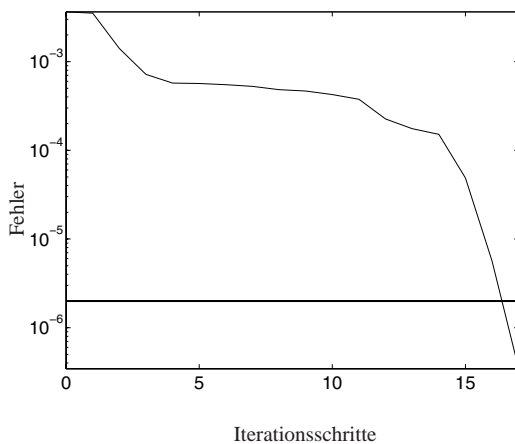
Für eine Scheibe, die sowohl durch einen Riß als auch durch ein ellipsenförmiges Loch geschädigt ist, ergeben sich die in den Abbildungen 5.23 und 5.24 dargestellten Optimierungsverläufe. In Abbildung 5.23 wird die Scheibe nur durch den Lastfall 1 belastet, in Abbildung 5.24 wird zusätzlich der zweite Lastfall verwendet. In beiden Fällen werden die Schädigungsparameter nach wenigen Iterationen identifiziert. Hier bewirkt die Hinzunahme eines zweiten Lastfalls keine Verbesserung der Identifikation oder der Verringerung der Rechenzeit. Die Optimierung dauert mit zwei Lastfällen sogar drei Iterationsschritte länger. Dieser Tatbestand resultiert daraus, daß die zu suchenden Schädigungen nicht übereinander liegen wie im Simulationsbeispiel nach Abbildung 5.22.



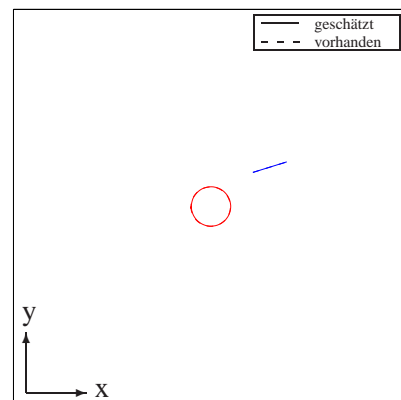
(a) Mittelpunktkoordinaten



(b) Durchmesser der Ellipse, Länge und Winkel des Risses

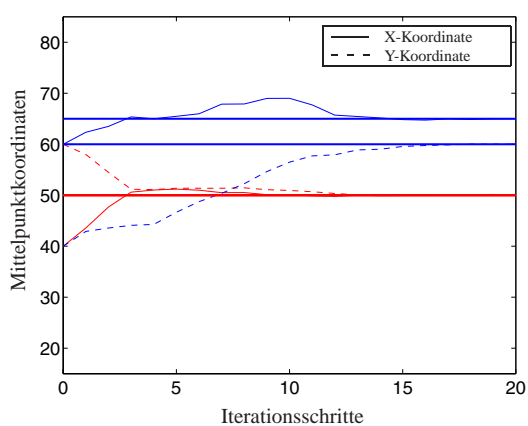


(c) Fehlerverlauf der Optimierung

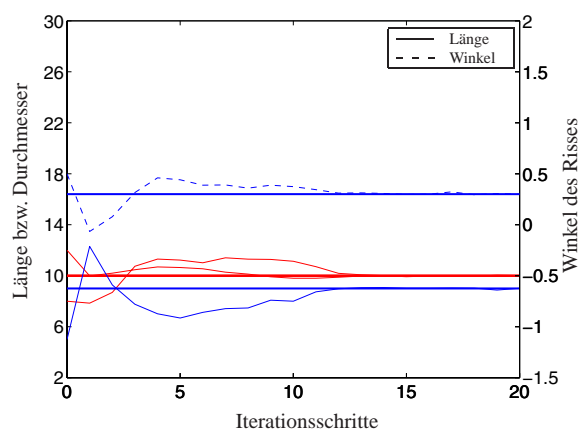


(d) Rißposition in der Scheibe

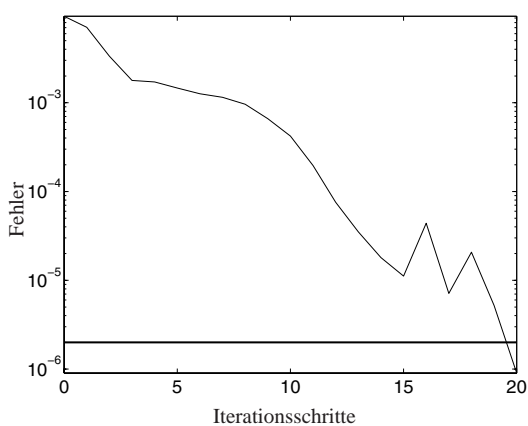
Abbildung 5.23: Identifikation einer Ellipse und eines Risses mit dem Kalman-Filter durch einen Lastfall



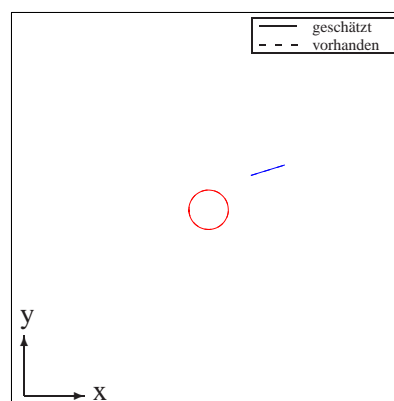
(a) Mittelpunktkoordinaten



(b) Durchmesser der Ellipse, Länge und Winkel des Risses



(c) Fehlerverlauf der Optimierung

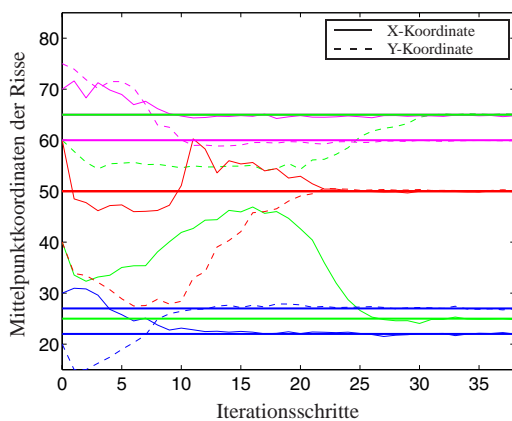


(d) Rißposition in der Scheibe

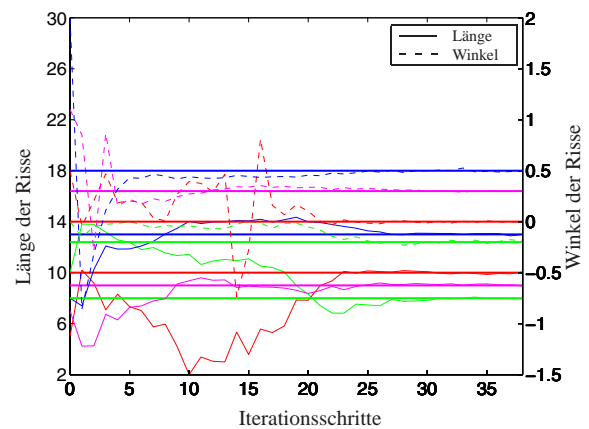
Abbildung 5.24: Identifikation einer Ellipse und eines Risses mit dem Kalman-Filter durch zwei Lastfälle

5.1.10 Identifikation von vier Rissen

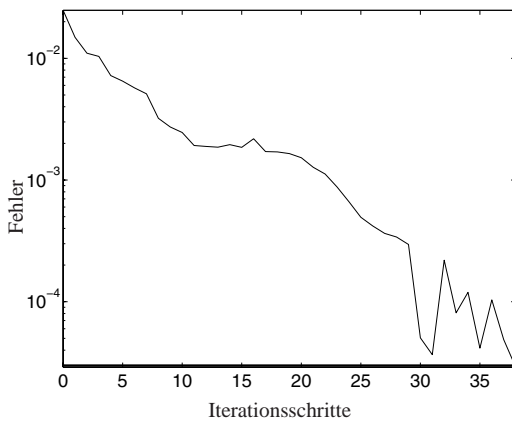
Die Optimierung einer durch vier Risse geschädigten Scheibe ist in Abbildung 5.25 gezeigt. Somit müssen 16 unbekannte Parameter durch das Kalman-Filter identifiziert werden. Beide Lastfälle werden zur Minimierung des Fehlerwertes verwendet. Der Verlauf des Fehlers ist in Abbildung 5.25 angegeben. Nach 37 Iterationen ist der Identifikationsvorgang erfolgreich abgeschlossen, und alle vier Risse konnten erkannt werden. Die relative Nähe der Startwerte zur gesuchten Rißposition wirkt sich bei der Optimierung positiv aus. Da die vorhandenen Risse ebenso wie die Startwerte über die ganze Scheibe verteilt liegen, sind Start- und Zielkoordinaten relativ dicht beieinander. Diese Tatsache vereinfacht den Optimierungsvorgang.



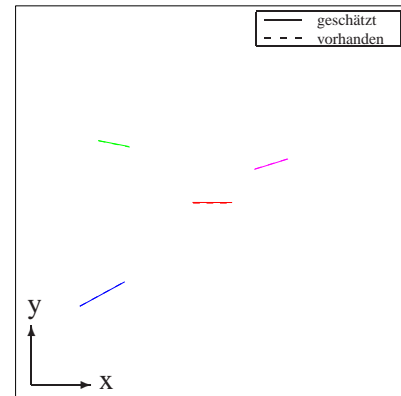
(a) Mittelpunktkoordinaten der Risse



(b) Länge und Winkel der Risse



(c) Fehlerverlauf der Optimierung



(d) Rißposition in der Scheibe

Abbildung 5.25: Identifikation von vier Rissen mit dem Kalman-Filter durch zwei Lastfälle

5.1.11 Schädigungsidentifikation in nicht rechteckigen Scheiben

In den vorherigen Abschnitten wurden nur Schädigungen in einer quadratischen Scheibe identifiziert. Nun soll die Rißidentifikation auf nichtkonvexe Geometrien wie L- oder U-Profile erweitert werden. Die Besonderheit im Optimierungsverlauf besteht nun darin, daß die zu suchenden Parameter nicht immer auf direktem Weg ermittelt werden können. Die direkte Verbindung zwischen den Startwerten und den zu suchenden Rißkoordinaten kann außerhalb der Geometrie verlaufen. Das Identifizieren von Defekten außerhalb der Struktur wird durch Nebenbedingungen unterbunden.

5.1.12 L-förmige Scheibe

Die untersuchten Belastungssituationen für eine L-förmige Scheibe sind in der Abbildung 5.26 dargestellt. Abbildung 5.27 zeigt die Abmessungen und die Diskretisierung des L-Profils. Das

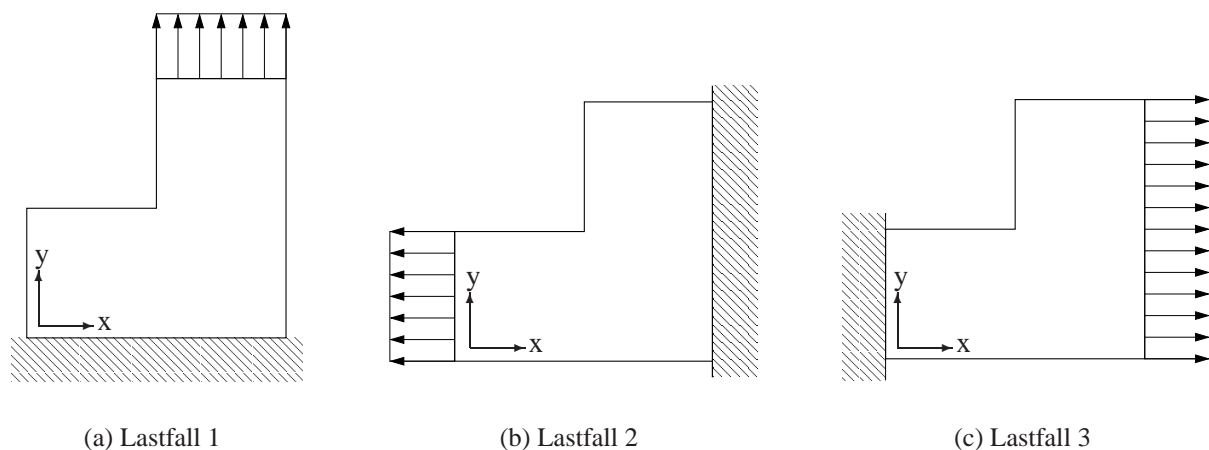


Abbildung 5.26: Die drei Belastungsfälle der L-förmigen Scheibe

Koordinatensystem befindet sich in der unteren linken Ecke des Profils. Die langen Seiten der Struktur sind mit zehn quadratischen Elementen vernetzt, die kurzen mit fünf. Für die rißförmigen Schädigungen wird die gleiche Diskretisierung verwendet wie in den vorherigen Untersuchungen.

In Abbildung 5.28 ist der Optimierungsverlauf unter Verwendung des Lastfalls 1 demonstriert. Der Startwert $\hat{x}_0^- = [30 \ 20 \ 8 \ 2, 0]$ liegt in dem unteren rechten Teil der Scheibe. Die Rißparameter sind nach 90 Iterationsschritten gut angenähert.

Liegt der zu suchende Riß im unteren linken Teil des L-Profils und die Startwerte befinden sich im oberen Teil der Scheibe nahe der Belastung, ergibt sich der in Abbildung 5.29 dargestellte Verlauf der Optimierung. Der Riß wird in der durch Lastfall 1 belasteten Scheibe nicht identifiziert. Es wird deutlich, daß die Lage des zu suchenden Risses unter der aufgebrachten Belastung keinen Einfluß auf die Verformungen der Scheibenaußenseite hat. Kennzeichnend hierfür ist der

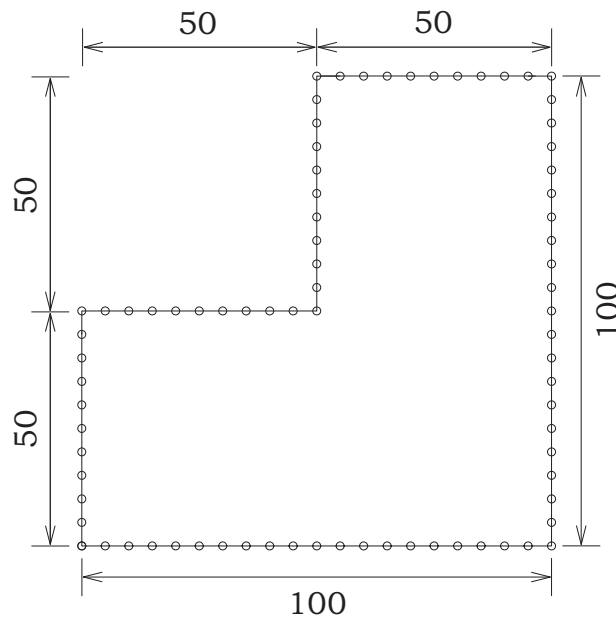


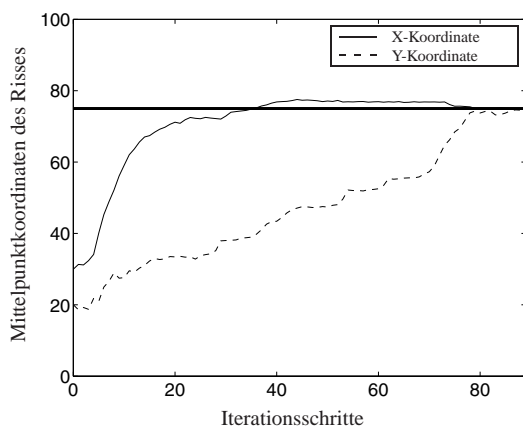
Abbildung 5.27: Diskretisierung und Abmessungen des L-Profiles

niedrige Fehlerwert in Abbildung 5.29c während der Optimierung. Das Kalman-Filter versucht die Rißlänge zu minimieren und stößt dabei an die untere vorgegebene Grenze.

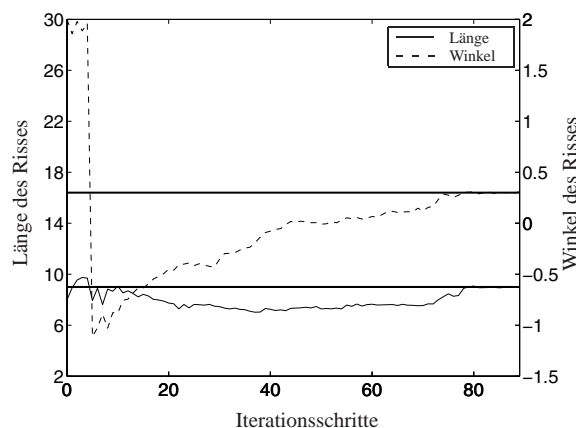
Ein weiteres Problem der Optimierung ist die nichtkonvexe Geometrie. Die Rißparameter können nicht auf direktem Wege erreicht werden.

Eine Anregung der Scheibe durch den zweiten Belastungsfall führt, da es sich hierbei um eine ähnliche Belastungssituation wie in Abbildung 5.28 handelt, zum Erfolg. Der Riß wird in dem Fall direkt dem durch die Zugbelastung erzeugten Kraftfluß ausgesetzt. Während Risse innerhalb rechteckiger Scheiben mit einem Lastfall unabhängig von deren tatsächlichen Lage gefunden werden, ist bei dieser L-förmigen Scheibe die Wahl des Lastfalls in Abhängigkeit von der Schädigungsposition entscheidend für die erfolgreiche Identifikation. Somit müssen immer mehrere Lastfälle entweder nacheinander oder gleichzeitig betrachtet werden. Ist die Identifikation mit einem Lastfall fehlgeschlagen, so ist es noch nicht auszuschließen, daß es sich um ein ungeschädigtes Bauteil handelt. Die Anwendung einer anderen oder zusätzlichen Belastung kann die Identifikation des Defektes ermöglichen.

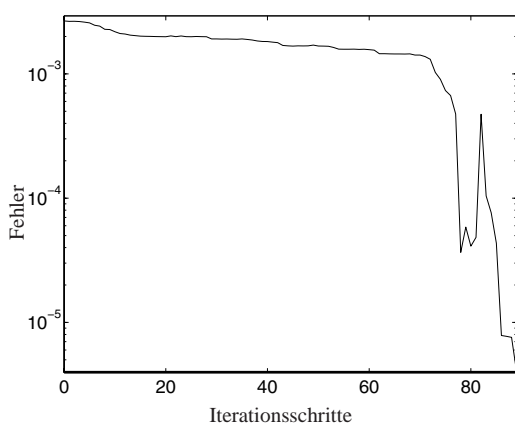
Auch mehrere Schädigungen in der L-förmigen Scheibe können identifiziert werden. In Abbildung 5.30 ist der Optimierungsverlauf für zwei Risse in einer durch die Lastfälle 1 und 3 belasteten Scheibe dargestellt. Die Rißparameter sind nach 140 Iterationen gefunden. Die Anwendung beider Lastfälle liefert wie auch in den Untersuchungen mit einer quadratischen Scheibe mehr Informationen, um die Rißgrößen korrekt zu erkennen. Die Identifikation dreier gleichverteilter Risse in der L-förmigen Scheibe zeigt Abbildung 5.31. Nach 45 Iterationen werden die drei Risse bei Anwendung von nur einer Belastung nach Lastfall 1 gut erkannt. Bei Vorgabe gleichverteilter Startwerte ist in diesem Fall die Identifikation mit nur einem Lastfall



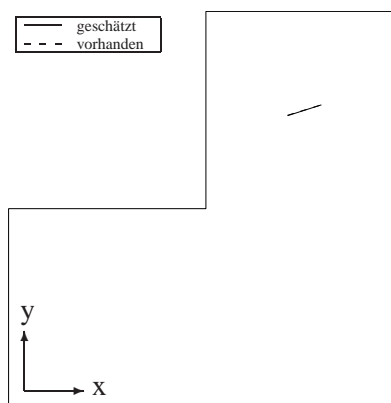
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



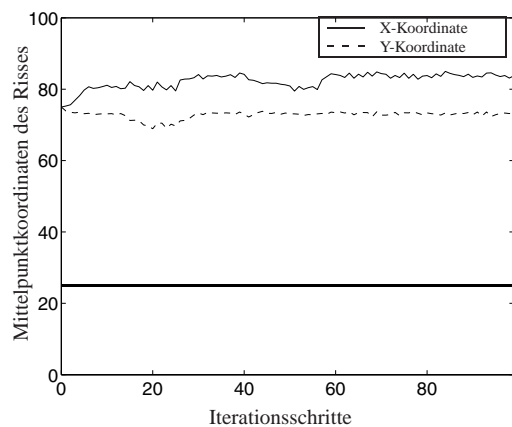
(c) Fehlerverlauf der Optimierung



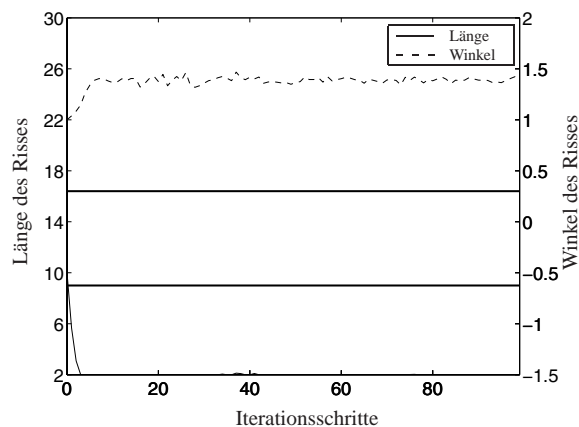
(d) Rißposition in der Scheibe

Abbildung 5.28: Identifikation eines Risses innerhalb einer L-förmigen Scheibe durch einen Lastfall mit einem Startwert von $[30 \ 20 \ 8 \ 2]$

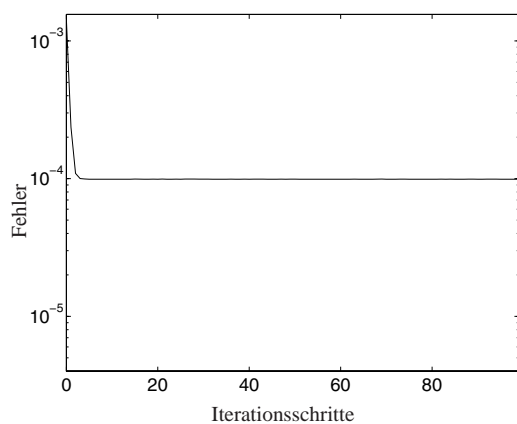
möglich, da die Differenz aus Anfangswerten und Endwerten des Risses klein ist und somit die Identifikation durch das Kalman-Filter erleichtert wird. Mit steigender Anzahl zu suchender gleichverteilter Risse steigt gleichzeitig die Wahrscheinlichkeit für einen kleinen Abstand der Mittelpunktkoordinaten gleichverteilter Startwerte zu den zu suchenden Rißparametern an. Dieser Sachverhalt wirkt sich positiv auf die Optimierung durch das Kalman-Filter aus.



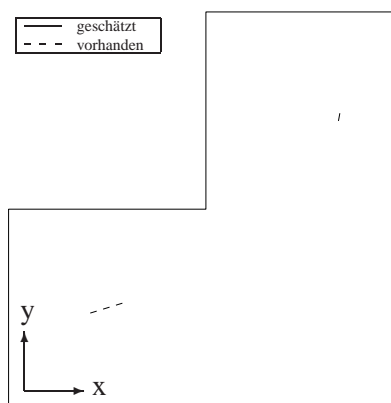
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

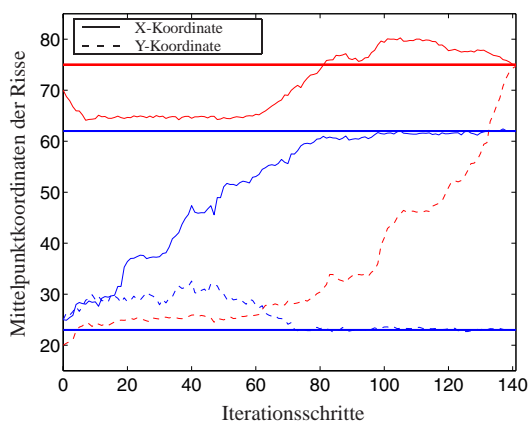


(c) Fehlerverlauf der Optimierung

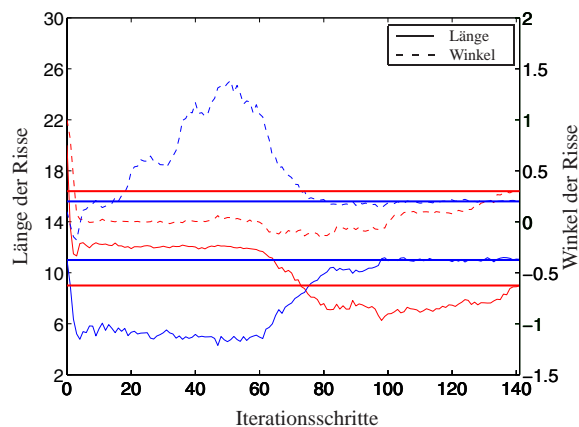


(d) Rißposition in der Scheibe

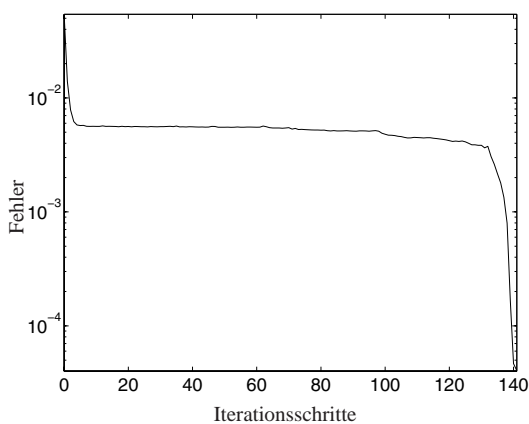
Abbildung 5.29: Identifikation eines Risses innerhalb einer L-förmigen Scheibe durch einen Lastfall mit einem Startwert von [75 75 10 1]



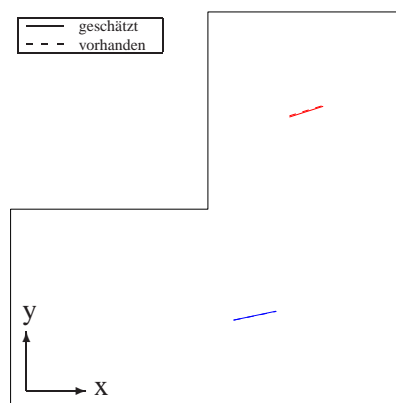
(a) Mittelpunktkoordinaten der Risse



(b) Länge und Winkel der Risse

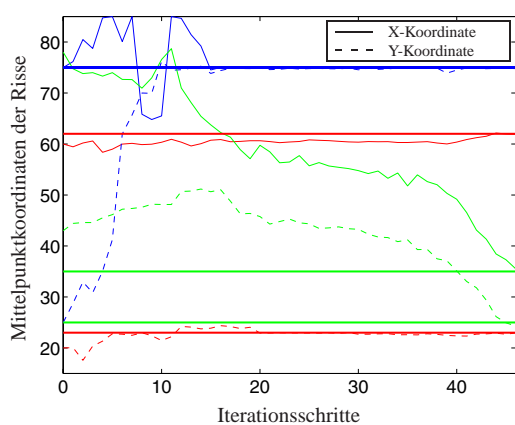


(c) Fehlerverlauf der Optimierung

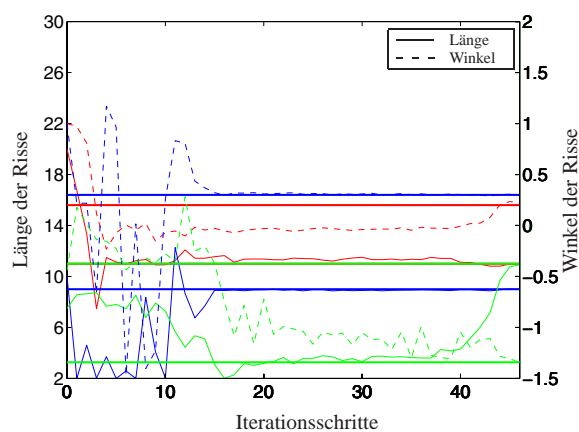


(d) Rißposition in der Scheibe

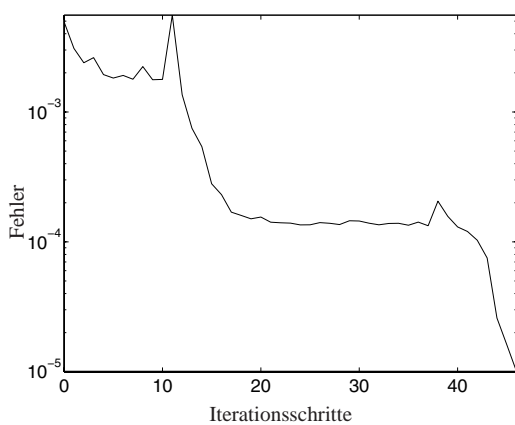
Abbildung 5.30: Identifikation zweier Risse innerhalb einer L-förmigen Scheibe durch die Lastfälle 1 und 3 mit den Startwerten von $[70 \ 20 \ 20 \ 1]$ und $[25 \ 25 \ 11 \ 0]$



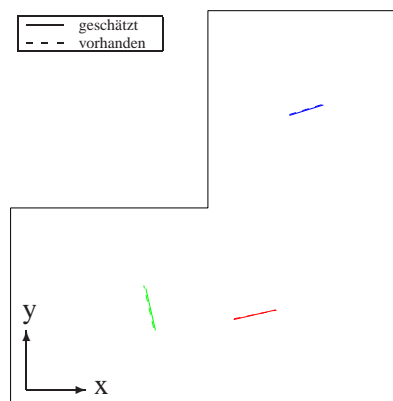
(a) Mittelpunktkoordinaten der Risse



(b) Länge und Winkel der Risse



(c) Fehlerverlauf der Optimierung



(d) Rißposition in der Scheibe

Abbildung 5.31: Identifikation von drei Rissen innerhalb einer L-förmigen Scheibe durch den Lastfall 1 mit den Startwerten von $[60 \ 20 \ 20 \ 1]$, $[75 \ 25 \ 10 \ 1]$ und $[78 \ 43 \ 7,5 \ 2,7]$

5.1.13 U-förmige Scheibe

Ein U-Profil soll in diesem Abschnitt untersucht werden. Die Geometrie wird durch sechs unterschiedliche Belastungssituationen nach Abbildung 5.32 angeregt. In Abbildung 5.33 sind die

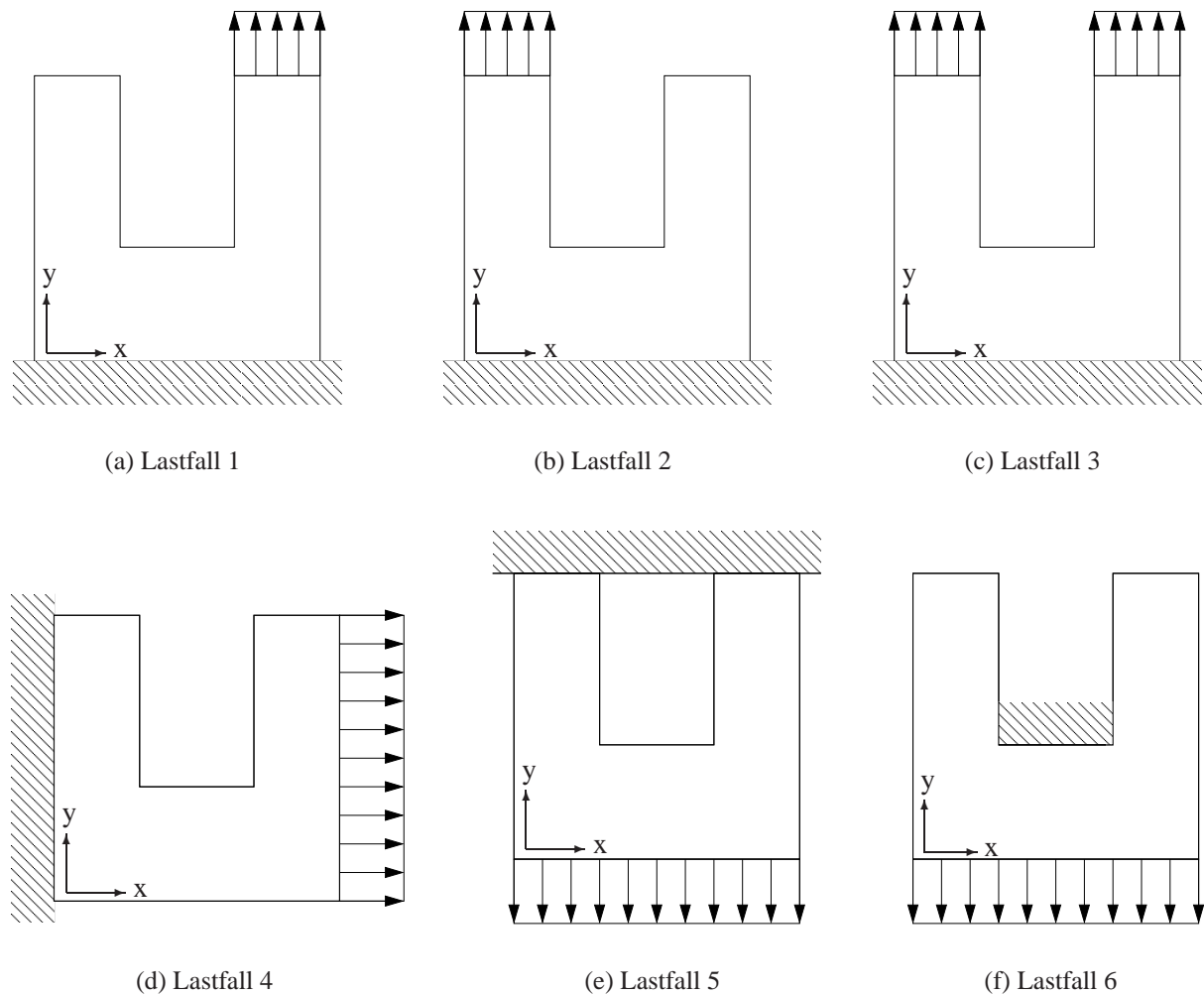


Abbildung 5.32: Die sechs Belastungsfälle für die U-förmige Scheibe

Diskretisierung des äußeren Randes und die Abmessungen der Struktur dargestellt.

Befindet sich eine rißförmige Schädigung in einem der seitlichen Flansche der U-Form, kann eines geeigneten Lastfalls identifiziert werden. Der in Abbildung 5.34 gezeigte Iterationsverlauf zeigt eine gute Konvergenz der Optimierung. Die Scheibe wird durch Lastfall 1, eine Belastung des geschädigten Flansches, angeregt. Der Startwert des Kalman-Filters wird im gegenüberliegenden nicht belasteten Teil der Scheibe vorgegeben. Der Iterationsverlauf der Mittelpunktkoordinaten nach Abbildung 5.34a zeigt ein Absinken der y -Koordinate mit gleichzeitigem Ansteigen der x -Koordinate. Im Anschluß ändert sich die x -Koordinate wenig und die y -Koordinate steigt bis zum Zielwert hin an. Dieser Verlauf beschreibt die „Wanderung“ des Risses innerhalb

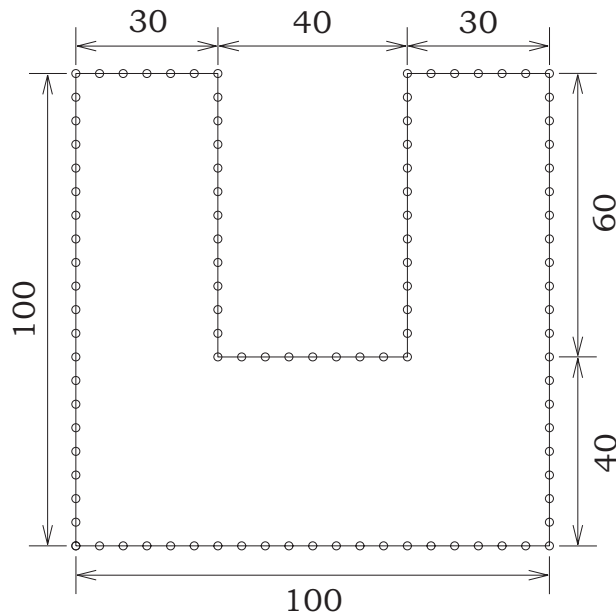


Abbildung 5.33: Diskretisierung und Abmessungen des U-Profiles

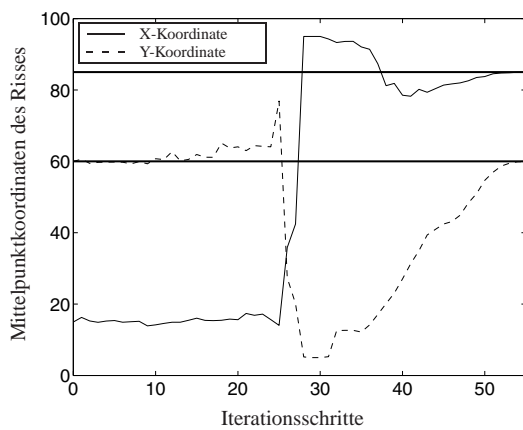
der U-Form. Während der Optimierung wird keine der Nebenbedingungen verletzt. Die immer wieder neu geschätzte Riposition befindet sich nie auerhalb der Scheibe.

Wird nur der ungeschädigte Flansch der Struktur belastet, ist eine Identifikation hingegen nicht möglich, wenn sich der Mittelpunkt des Startwertes im ungeschädigten Teil der Struktur befindet. Dies ist in Abbildung 5.35 gezeigt, in der die Scheibe durch den Lastfall 2 angeregt wird. Während des gesamten Optimierungsverlaufes verbleiben die Mittelpunktkoordinaten in der Nähe ihrer Startwerte. Das Kalman-Filter reduziert die Rilänge auf die untere vorgegebene Grenze und versucht somit den Einflu des Risses auf die Verformungen zu reduzieren. Ein derartiger Optimierungsverlauf würde darauf hindeuten, da es sich um ein ungeschädigtes Bauteil handelt.

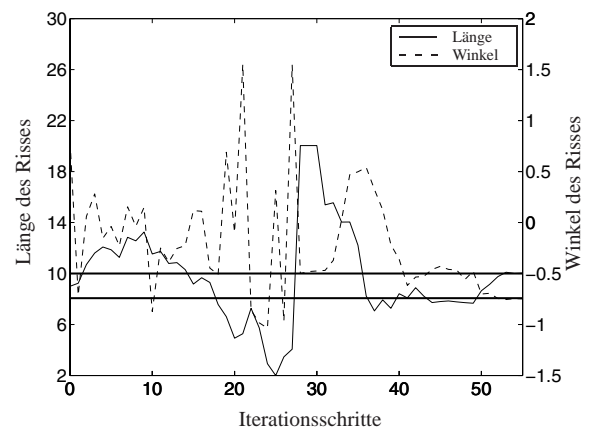
Wenn der geschädigte Flansch mit einer Belastung nach Lastfall 4 angeregt wird, milingt die Identifikation ebenfalls. Der Optimierungsverlauf ist in Abbildung 5.36 gezeigt. Die Startwerte ändern ihre Werte während der Optimierung ebenfalls nicht wesentlich. Das Kalman-Filter versucht auch hier, den Einflu des Risses innerhalb der Struktur an der prognostizierenden Stelle zu minimieren, in dem die Rilänge während der Optimierung gegen die untere Grenze der Nebenbedingung reduziert wird.

Die Belastungsform mit der die Identifikation eines Risses im Flansch der Struktur wesentlich zuverlässiger funktioniert, ist der Lastfall 3. In diesem werden beide Flansche gleichzeitig belastet. Abbildung 5.37 zeigt den Iterationsverlauf für diesen Fall. In Abbildung 5.37a kann die „Wanderung“ des Risses durch den Steg des U-Profiles nachvollzogen werden.

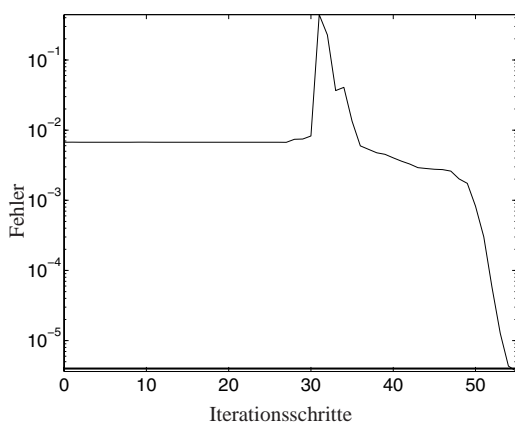
Weitere Tests haben ergeben, da die Identifizierung eines Risses im Steg der U-förmigen Scheibe durch die Lastfälle 1 bis 5 nicht möglich ist, wenn die Mittelpunktkoordinaten der Startwerte



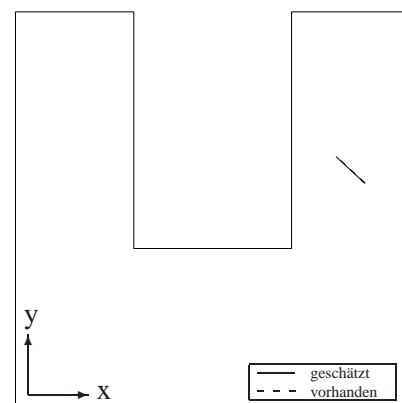
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

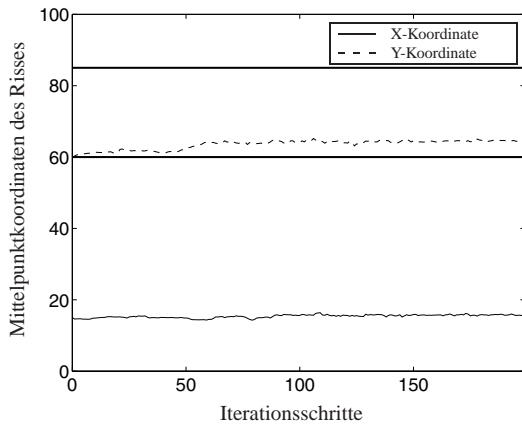


(c) Fehlerverlauf der Optimierung

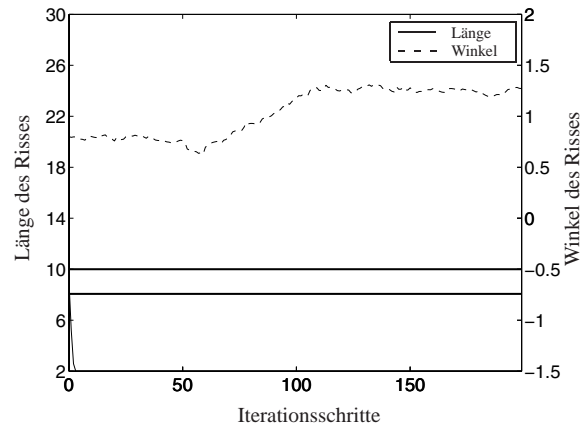


(d) Rißposition in der Scheibe

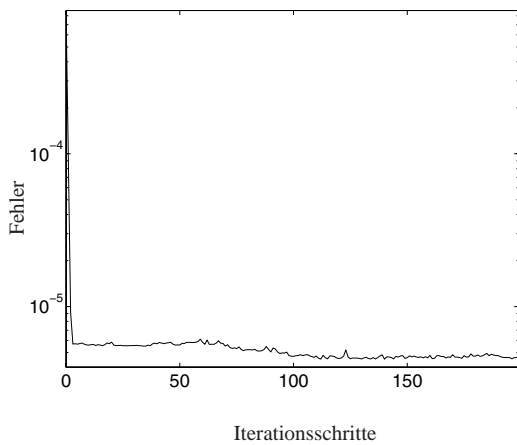
Abbildung 5.34: Identifikation eines Risses innerhalb einer U-förmigen Scheibe durch den Lastfall mit den Startwerten von [15 60 9 0,8]



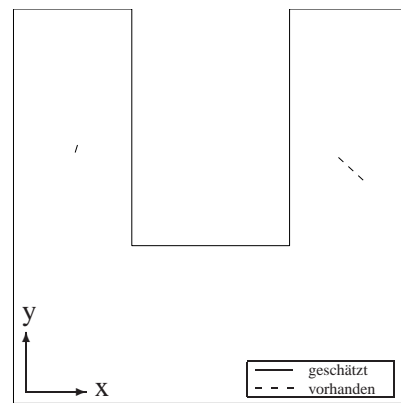
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



(c) Fehlerverlauf der Optimierung

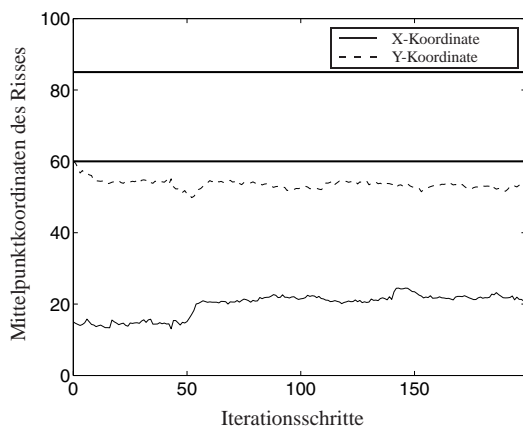


(d) Rißposition in der Scheibe

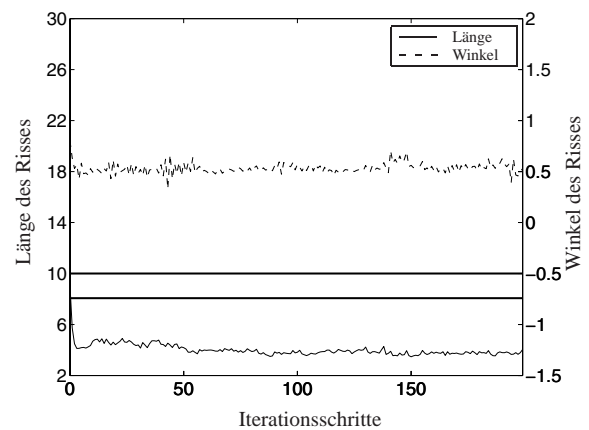
Abbildung 5.35: Identifikation eines Risses innerhalb einer U-förmigen Scheibe durch den Lastfall 2 mit den Startwerten von [15 60 9 0,8]

in einem der Flansche liegen. Die Mittelpunkte der geschätzten Rißparameter verbleiben im Flansch.

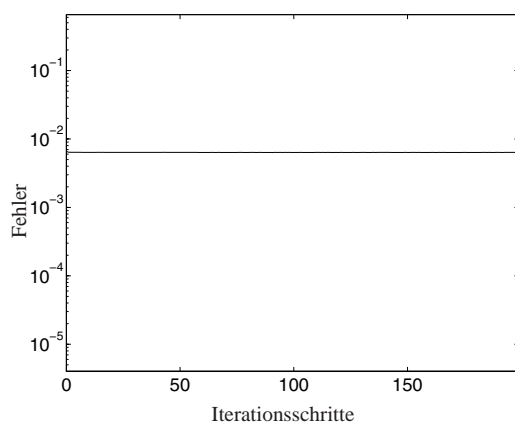
Die Identifikation gelingt nur mit dem im Labor schlecht nachstellbaren Belastungsfall 6. Die Parameter sind allerdings nicht sehr gut geschätzt, wie Abbildung 5.39 zeigt. Die Länge des Risses wird am schlechtesten wiedergegeben. Für die Identifikation einer Fehlstelle innerhalb eines U-Profils muß ein geeigneter Lastfall verwendet werden. Die Identifizierung eines Risses im Steg des U-Profils ist nur unter einer im Labor schlecht nachstellbaren Belastung möglich.



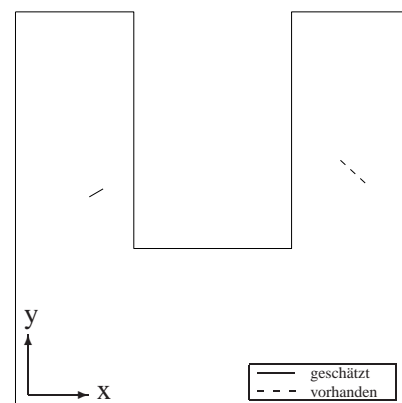
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

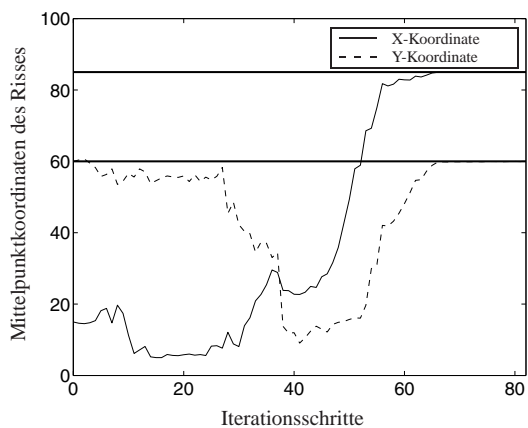


(c) Fehlerverlauf der Optimierung

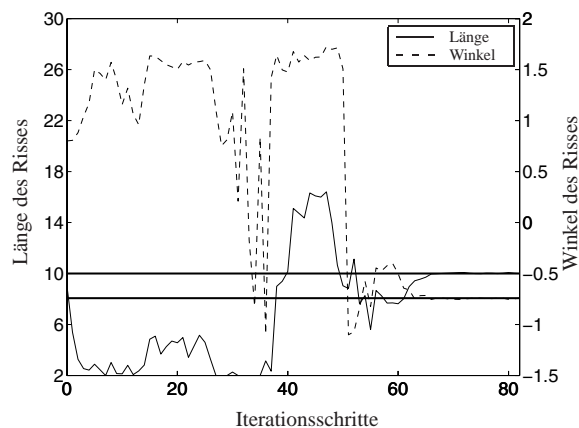


(d) Rißposition in der Scheibe

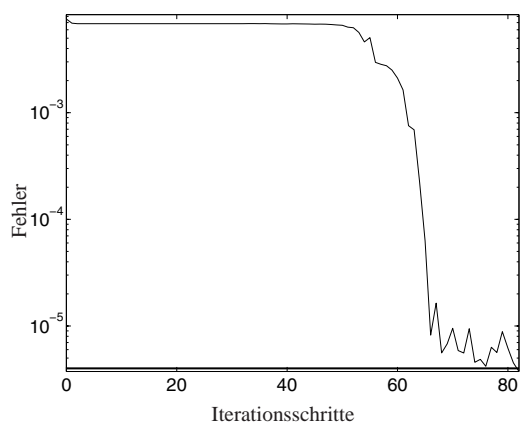
Abbildung 5.36: Identifikation eines Risses innerhalb einer U-förmigen Scheibe durch den Lastfall 4 mit den Startwerten von [15 60 9 0,8]



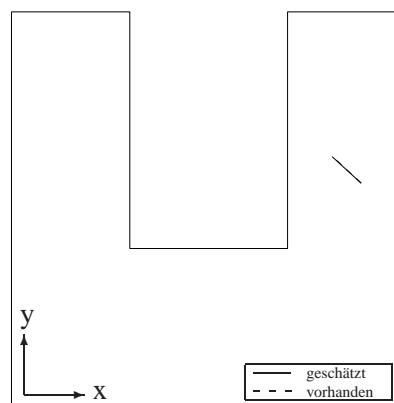
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

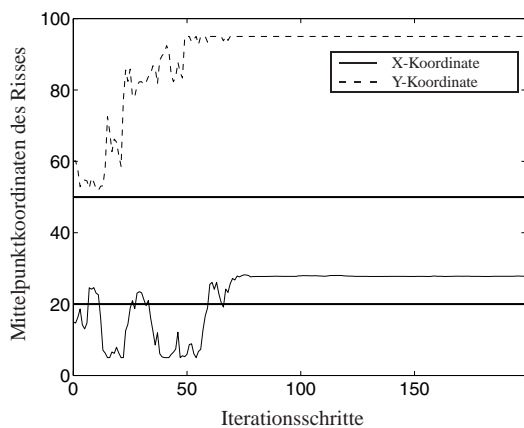


(c) Fehlerverlauf der Optimierung

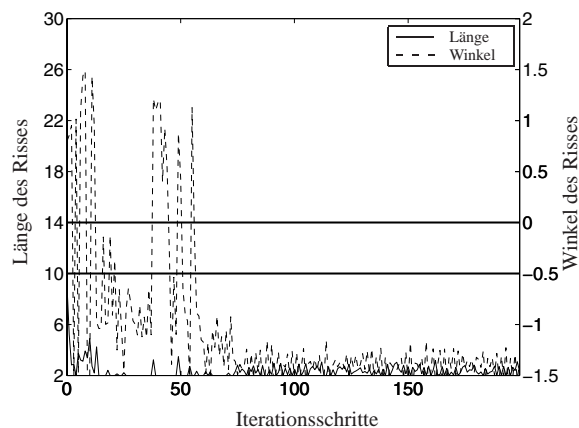


(d) Rißposition in der Scheibe

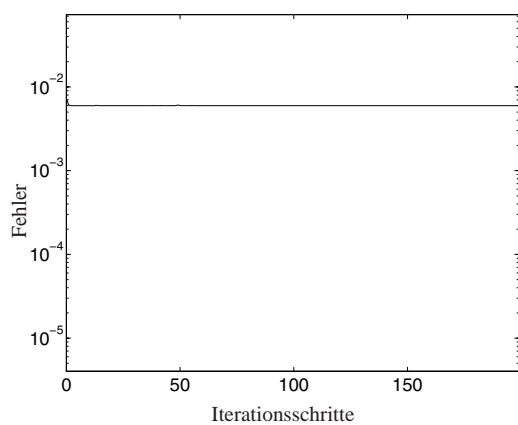
Abbildung 5.37: Identifikation eines Risses innerhalb einer U-förmigen Scheibe durch den Lastfall 3 mit den Startwerten von [15 60 9 0,8]



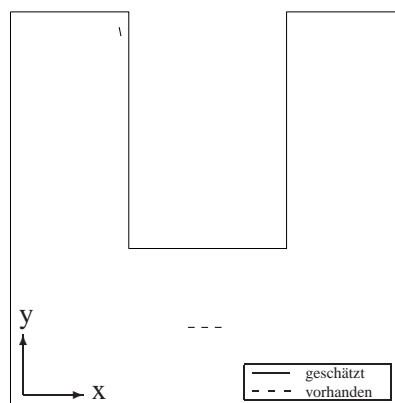
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses

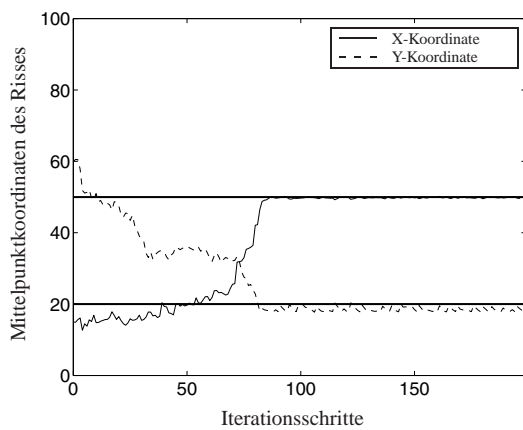


(c) Fehlerverlauf der Optimierung

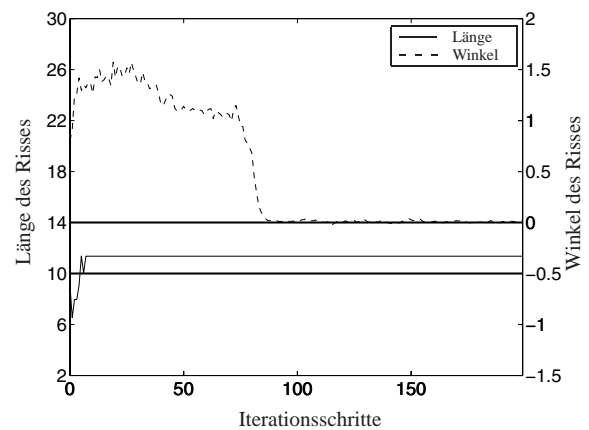


(d) Rißposition in der Scheibe

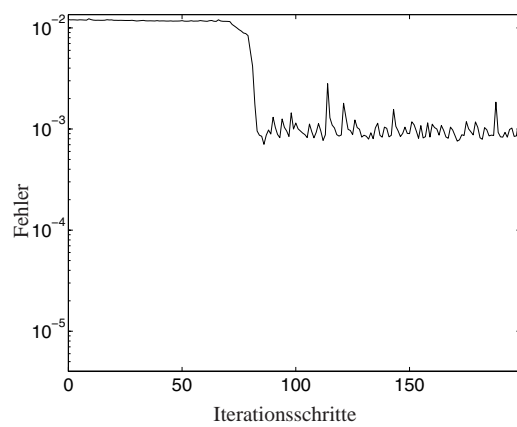
Abbildung 5.38: Identifikation eines Risses innerhalb einer U-förmigen Scheibe durch den Lastfall 5 mit den Startwerten von [15 60 9 0,8]



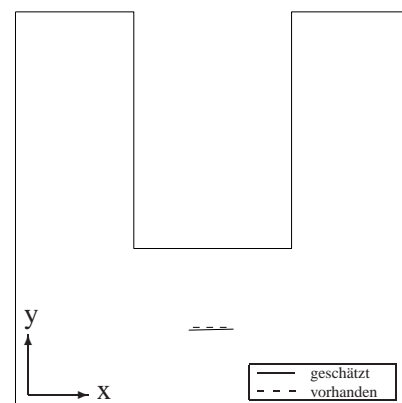
(a) Mittelpunktkoordinaten des Risses



(b) Länge und Winkel des Risses



(c) Fehlerverlauf der Optimierung



(d) Rißposition in der Scheibe

Abbildung 5.39: Identifikation eines Risses innerhalb einer U-förmigen Scheibe durch den Lastfall 6 mit den Startwerten von [15 60 9 0,8]

5.2 Identifikation mit neuronalen Netzen

Die Schadenserkennung mit neuronalen Netzen wurde in Kapitel 4.3 eingehend für statisch belastete Geometrien untersucht. Der Erfolg der Identifikation hing weitgehend von der Wahl des Optimierungsalgorithmus ab, der zum Trainieren des Netzes erforderlich ist. Bei der Identifikation im Zeitbereich wird ausschließlich der Levenberg-Marquardt Algorithmus eingesetzt, da diese Strategie die besten Ergebnisse bei Untersuchungen in der Statik lieferte. Die Untersuchungen beschränken sich auf quadratische Scheiben, und es werden ausschließlich die Mittelpunktkoordinaten der Fehlstelle prognostiziert. Die Diskretisierung des Scheibenrandes geht aus Abbildung 5.1 hervor. Es werden 10 quadratische Elemente je Scheibenseite verwendet. Eine rißförmige Fehlstelle wird mit nur zwei quadratischen Elementen versehen. Ein lochförmiger Defekt wird hier mit fünf quadratischen Elementen diskretisiert.

Die Schadensidentifikation in einer dynamisch belasteten Scheibe ähnelt im Prinzip der statisch belasteten Geometrie abgesehen von der Auswahl der Meßpunkte. Zudem gibt es nicht nur einen Meßwert pro Meßpunkt, sondern durch den Einfluß des zeitlichen Verhaltens entsteht je betrachteten Zeitpunkt ein Meßwert. Es stehen somit für das neuronale Netz sehr viele Verformungswerte der Außenseite der Scheibe als Eingangsgrößen zur Verfügung. Ein Netz mit einer sehr großen Eingangsschicht benötigt, um nicht übertrainiert zu werden, sehr viele Lernbeispiele. Die Erhöhung der Anzahl der Lernbeispiele ist nicht ohne weiteres möglich, da die Rechenzeit stark ansteigt. Eine Auswahl der Meßwerte muß durchgeführt werden.

Hierzu werden in einem ersten Schritt mehrere Zeitbereichsrechnungen mit der Randelementmethode durchgeführt, um den zeitlichen Verlauf der Verformungen aller Diskretisierungspunkte aufzunehmen. Nacheinander wird der Ort der Schädigung innerhalb der Scheibe verändert, um so die Punkte zu erfassen, deren Verformungswerte geeignet zum Trainieren des Netzes erscheinen. Im einzelnen werden für lochförmige Schädigungen 81 unterschiedliche Simulationen durchgeführt. Beginnend mit einer Fehlstelle bei den Koordinaten $x_c = 10$ und $y_c = 10$ werden diese in Schritten von jeweils zehn je Koordinatenrichtung erhöht. Der letzte Datensatz beinhaltet alle zeitlichen Verformungswerte der gesamten Außenseite der Scheibe für eine Schädigung an der Position $x_c = 90$ und $y_c = 90$. Der Lochdurchmesser wird während der Berechnungen mit $d = 10$ konstant gehalten.

Im Gegensatz zur Simulation lochförmiger Fehlstellen wird bei der Simulation rißförmiger Fehlstellen die benötigte Rechenzeit ansteigen, da der mögliche Kontakt der Rißflächen berücksichtigt werden muß. Aus diesem Grunde werden nur 64 Datensätze generiert. Hier wird der Mittelpunkt des Risses beginnend mit den Koordinaten $x_c = 20$ und $y_c = 20$ in Zehnerschritten bis zum letzten betrachteten Punkt $x_c = 80$ und $y_c = 80$ innerhalb der quadratischen Scheibe variiert. Dabei wird die Rißlänge mit $l_c = 10$ und der Neigungswinkel des Risses zu $\alpha = 0$ fest vorgegeben. Die Scheibenoberseite wird einmal auf Druck und zum anderen auf Zug nach Abbildung 5.2 belastet, während die Scheibenunterseite fest eingespannt ist. In Abbildung 5.40 ist der zeitliche Verschiebungsverlauf des Mittelpunktes der zugbelasteten Scheibenoberseite für alle verwendeten 64 Schadenspositionen eines Risses dargestellt. Für alle Schädigungspositionen ergibt sich in den ersten zeitlichen Simulationsschritten ein identischer Verlauf. Ab etwa vier Simulationsschritten weist die Verschiebung für einige Graphen unterschiedliche Werte

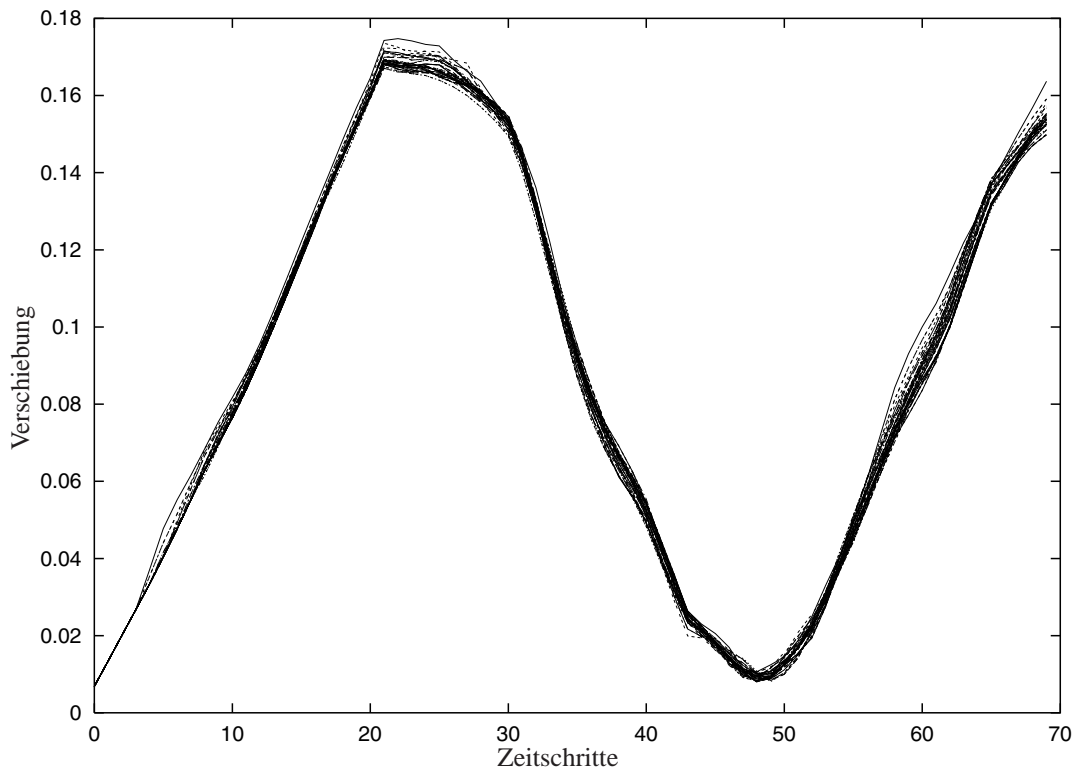


Abbildung 5.40: Zeitlicher Verformungsverlauf ermittelt am Punkt A unter Zugbelastung

auf. Zu diesem Zeitpunkt sind die ersten Wellenreflektionen, die durch eine Schädigung hervorgerufen werden, zu erkennen. Eine Schädigung, die sehr nahe an der Scheibenoberseite positioniert ist, zeichnet sich durch eine sehr frühe Abweichung des zeitlichen Verformungsverhaltens gegenüber dem des ungeschädigten Originals aus. Der zurückzulegende Weg der Welle von Meßpunkt der Scheibenoberseite zum Ort der Schädigung ist wesentlich geringer, als der Weg, den die Wellenfront zurücklegen muß, wenn sich die Schädigung nahe der Einspannstelle befindet. Der zeitliche Verformungsverlauf des gleichen Meßpunktes erzeugt durch eine druckbelastete Scheibe ist in Abbildung 5.41 für alle 64 unterschiedliche Schadenspositionen dargestellt. Die Randbedingungen für diese Simulation sind bis auf die umgekehrte Richtung der Belastung gleich gehalten. Auffallend ist der nahezu identische Verlauf aller berechneten Verformungsverläufe des betrachteten Meßpunktes. Dies läßt darauf schließen, daß die Lage des Risses einen geringen Einfluß auf das Systemverhalten der druckbelasteten Scheibe hat. Die Identifizierung einer möglichen Schädigung wird somit erschwert.

Zur Vorbereitung des Lerndatensatzes wird für jeden Knoten der Scheibe dessen zeitliches Verformungsverhalten simuliert. Um die Auswahl an Meßpunkten und Meßzeitpunkten zum Aufbau des neuronalen Netzes einzugrenzen, werden Meßpunkte ausgewählt, deren zeitlicher Verschiebungsverlauf in Abhängigkeit der Schädigungsposition zu bestimmten Zeiten stärker voneinander abweichen. Es ist leicht aus den theoretischen Grundlagen nach Kapitel 3.4 eines neuronalen Netzes ersichtlich, daß stark streuende Werte als Eingangsgrößen die Leistungsfähigkeit des Netzes steigern können. Die gemessenen Verformungen gleich zu Beginn der Si-

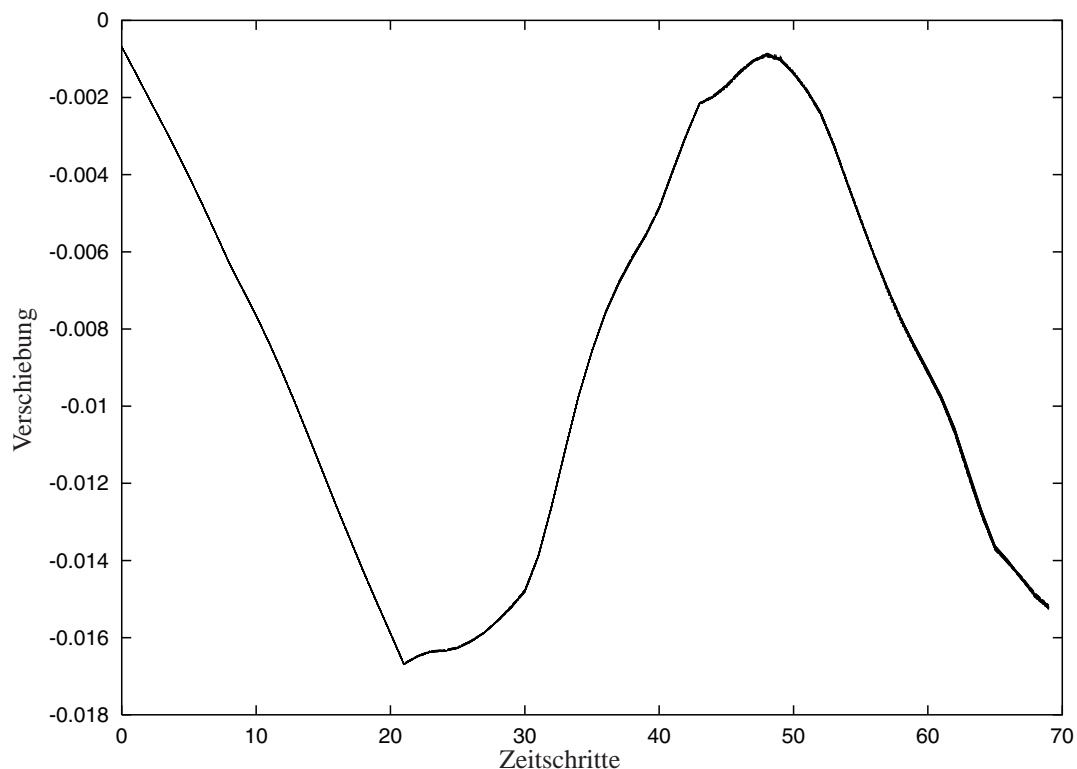


Abbildung 5.41: Zeitlicher Verformungsverlauf ermittelt am Punkt A unter Druckbelastung

mulation enthalten keine Informationen, durch die eine erfolgreiche Identifikation durchgeführt werden kann. Es muß die Reflektion der Welle abgewartet werden. Die Meßpunkte werden gleichmäßig am Rand verteilt angenommen und entsprechende Zeitpunkte herausgesucht. Beispielfhaft werden die Verformungen in Belastungsrichtung des 22., 30., 40., und 49. Zeitschrittes als Eingangsgrößen für das neuronale Netz zur Verfügung gestellt. Bei den in diesem Abschnitt vorgestellten Ergebnissen werden 18 Meßpunkte ausgewählt und insgesamt 206 Meßwerte, gemessen in horizontaler als auch in vertikaler Richtung, als Eingangsgrößen für das neuronale Netz ausgesucht. Weitere Untersuchungen haben ergeben, daß die Wahl der Meßpunkte nicht so entscheidend für die Schadenserkennung ist wie die Auswahl der Zeitpunkte, an denen gemessen wird. Eine gleichverteilte Auswahl an Meßpunkten am Scheibenrand ist in diesem Fall ausreichend.

In Abbildung 5.42 ist das Identifikationsergebnis einer zugbelasteten Scheibe dargestellt. Die Form der Schädigung entspricht der eines Loches. Nach einem Trainingsvorgang von 2.358 Iterationen kann der mittlere Lernfehler (MSE) des Netzes die vorgegebene Grenze von 0,05 unterschreiten. Das Ergebnis des Lernvorganges kann Abbildung 5.42a entnommen werden. Die Mittelpunktkoordinaten der zu suchenden Schädigungen sind mit 'o' dargestellt, während die durch das Netz prognostizierten mit '+' markiert sind. Das Training kann mit guten Ergebnissen, wie in Abbildung 5.42a ersichtlich, beendet werden. Der Testdatensatz nach Abbildung 5.42b wird ebenfalls mit gutem Resultat wiedergegeben. Somit kann das neuronale Netz erfolgreich für die Schadensidentifikation in einer zugbelasteten Scheibe eingesetzt werden.

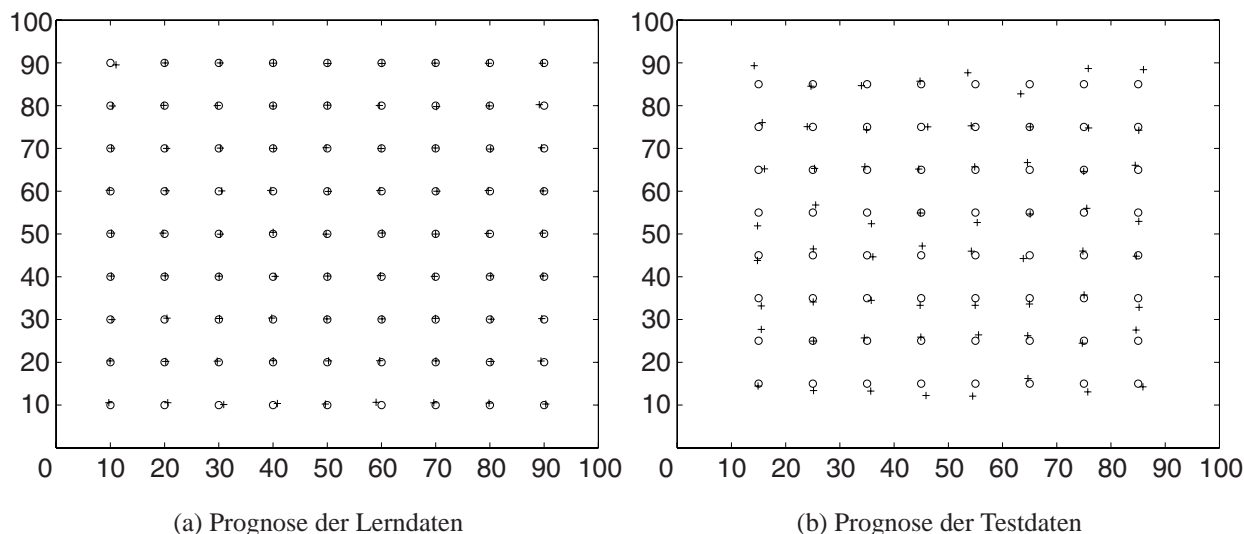


Abbildung 5.42: Prognose der Lern- und Testdaten für eine zugbelastete Scheibe mit lochförmiger Schädigung
 o vorgegebene Position + identifizierte Position

In einem nächsten Beispiel wird die Möglichkeit der Rißerkennung in einer druckbelasteten Scheibe betrachtet. Zum einen wird ein möglicher Kontakt der Rißflächen berücksichtigt, zum anderen wird der Riß ohne Kontakt simuliert. Letzteres entspricht nicht der physikalischen Interpretation. Ein überlappen der Rißflächen wird nicht ausgeschlossen und die Ergebnisse sollen nur zu Vergleichszwecken verwendet werden. In Abbildung 5.43 sind die Ergebnisse der Rißidentifikation mit und ohne Kontakt gegenübergestellt. Abbildung 5.43a verdeutlicht das Ergebnis des Trainingsvorgangs des neuronalen Netzes, wenn auf die Berücksichtigung des Kontaktes verzichtet wird. Der Lernvorgang kann erfolgreich abgeschlossen werden, und nach Abbildung 5.43b werden die Testdaten mit guter Genauigkeit erkannt. Wird die gleiche Schadenslokalisierung mit den Kontakteigenschaften des Risses durchgeführt, kann kein erfolgreicher Lernvorgang durchgeführt werden. Nach 2.500 Iterationen wird das Training beendet. Die Leistungsfähigkeit des Netzes zu diesem Zeitpunkt ist in den Abbildungen 5.43c und 5.43d dargestellt. Da das Training nicht mit genügender Genauigkeit abgeschlossen werden kann, sind auch die Prognosen des Testdatensatzes unbrauchbar.

Die Untersuchungen haben gezeigt, daß das neuronale Netz durchaus in der Lage ist, eine Schadensidentifikation erfolgreich durchzuführen. Auf Grund hoher Rechenzeiten wird sich hier im Gegensatz zur Statik auf die Identifikation der Schadensmittelpunkte beschränkt. In druckbelasteten Strukturen unter Berücksichtigung des Kontaktes kann die Lokalisierung der Fehlstelle nicht durchgeführt werden.

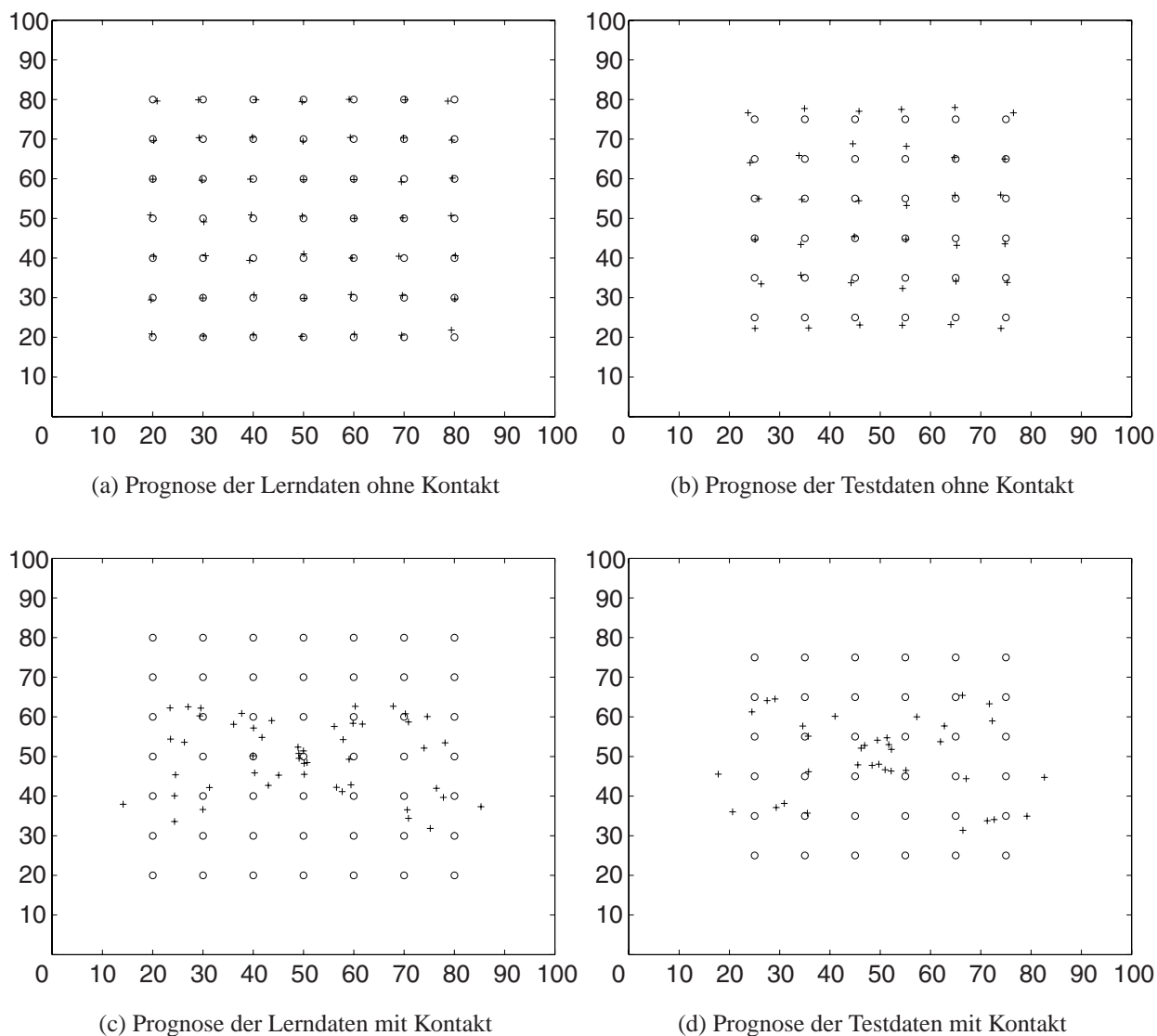


Abbildung 5.43: Prognose der Lern- und Testdaten für eine druckbelastete Scheibe mit rißförmiger Schädigung mit und ohne Berücksichtigung des Kontaktes der Rißflächen

o vorgegebene Position + identifizierte Position

6 Diskussion der Ergebnisse

Die numerischen Untersuchungen haben gezeigt, daß Fehlstellen innerhalb einfacher Geometrien, z.B. in einer Scheibe oder in einem Würfel, mit Optimierungsverfahren erkannt werden können. Es wurden die fünf folgenden Optimierungsstrategien

- Sequential Quadratic Programming (SQP)-Verfahren,
- linearisierter Kalman-Filter,
- Genetischer Algorithmus,
- Evolutionsstrategie,
- Neuronales Netz

hinsichtlich Ihrer Einsetzbarkeit zum Erkennen von Fehlstellen untersucht. Es stellte sich dabei heraus, daß es zwar mit allen Verfahren möglich ist, Schädigungen zu lokalisieren, jedes Verfahren jedoch unterschiedliche Vor- und Nachteile aufweist, die im Folgenden dargelegt werden sollen.

Die SQP-Verfahren gelten als leistungsfähige Optimierungsverfahren der klassischen, mathematischen Optimierung. Die Bestimmung eines Optimums erfolgt durch Lösung der Karush-Kuhn-Tucker-Bedingungen, die durch die Problemstellung festgelegt sind. Sie sind insbesondere für Optimierungsprobleme geeignet, deren Zielfunktion einen nichtlinearen Verlauf haben. Es ergibt sich die Notwendigkeit, die Jacobimatrix in das Verfahren miteinzubeziehen, wobei zu deren Berechnung eine Annäherung der Differentialquotienten durch zentrale Differenzenquotienten vorgenommen wird. Für eine erfolgreiche Optimierung werden, wie auch bei anderen Verfahren der mathematischen Optimierung, geeignete Startwerte benötigt.

Es konnte gezeigt werden, daß einfache Rißidentifizierungsverfahren, in denen nur die Lage der Fehlstelle zu bestimmen sind, auch bei sehr ungünstiger Wahl der Startwerte ohne weitere Probleme lösbar sind. Bei Ausweitung der Optimierung auf mehrere Unbekannte gewinnt die Wahl der Startwerte an Einfluß auf den Verlauf der Optimierung. Bei willkürlicher Wahl der Startwerte gelang es nicht, die zwei maßgebenden Parameter einer elliptischen Fehlstelle, die Lage und Größe der Halbachsen, innerhalb einer quadratischen Scheibe zu identifizieren. Nur bei einer Aufgabenstellung, bei der eine Vermutung über die Lage und über das Ausmaß des Defektes vorliegt, kann davon ausgegangen werden, daß das SQP-Verfahren gegen die richtige

Lösung konvergiert. Weiterhin ist festzustellen, daß das Verfahren erheblich mehr Iterationsschritte benötigt als das alternativ eingesetzte Kalman-Filter. Ausführliche Tests mit identischer Optimierungsaufgabe belegen dieses Ergebnis. Aufgrund dessen wurde zur grundsätzlich numerisch aufwendigeren Rißerkennung bei dynamischen Belastungen nur das Kalman-Filter weiter verwendet.

Das Kalman-Filter wurde ursprünglich entwickelt, um Schätzwerte interner nicht meßbarer Zustandsgrößen eines dynamischen Systems an Hand seiner Ein- und Ausgangsgrößen zu erhalten. Durch gezielte Umformulierung kann es als Optimierungsverfahren und damit zur Fehlstellen-erkennung eingesetzt werden. Das Kalman-Filter berücksichtigt bei den Berechnungen eventuell auftretende Störgrößen und verrauschte Signale, in dem die beschreibenden Gleichungen stochastische Größen enthalten. Bei Vergleich mit experimentell gewonnenen Daten besteht eine Möglichkeit, Meßfehler in Form dieser statistischen Größen in das Modell des Kalman-Filters direkt zu integrieren. Wie beim SQP-Verfahren benötigt das Kalman-Filter Gradienteninformationen des Optimierungsproblems. Diese werden auch hier mit dem zentralen Differenzenquotienten angenähert. Ebenso müssen dem Algorithmus Startwerte als Eingangswerte zur Verfügung gestellt werden.

Die Untersuchungen ergaben, daß das Kalman-Filter sehr gut geeignet ist, Fehlstellen auf Basis von Daten aus statischen Belastungsfällen und aus dem Zeitbereich zu erkennen. Die Simulationen in der Statik zeigen, daß das Kalman-Filter das schnellste iterative Verfahren ist, das in dieser Arbeit getestet wurde. Aus diesem Grund wurde die Anwendung im Zeitbereich weiter verfolgt. Untersuchungen im Dreidimensionalen ergaben allerdings, daß das Filter mit der Nichtkonvexität der Optimierung Probleme hatte. In einigen Problemstellungen konnte das Kalman-Filter nur ein lokales Minimum errechnen. Somit schlug die Schadensidentifikation fehl. Die vielen einzelnen Testergebnisse wurden nicht in dieser Arbeit gezeigt. Der Einfluß der Wahl der Startwerte ist nicht so gravierend wie beim SQP-Verfahren. Studien im Zeitbereich zeigen allerdings, daß bei schlecht gewählten Startwerten nur lokale Minima gefunden werden. Der Einfluß der unterschiedlichen Kovarianzmatrizen auf den Verlauf der Optimierung wurde ebenfalls analysiert. Das Kalman-Filter bietet als einziges der in dieser Arbeit getesteten Optimierungsverfahren die Möglichkeit an, auch verrauschte Meßsignale oder andere Störgrößen in die Berechnung der Schadensidentifikation mit einzubeziehen.

Ebenso wurde der Einfluß fehlerhafter Eingangsdaten auf den Erfolg der Identifizierung untersucht. Die aus numerischen Simulationen exakt gewonnenen Ergebnisse wurden mit einem prozentualen Fehler behaftet und als Daten für die Identifikation eingesetzt. Wie zu erwarten, nahm die Genauigkeit der Ergebnisse im Vergleich zu den exakten Daten ab, blieb jedoch im für die Schadenserkennung akzeptablem Bereich.

Der genetische Algorithmus und die Evolutionsstrategie sind der Theorie nach die einzigen hier getesteten Optimierungsverfahren, die in der Lage sind, lokale Minima zu überwinden und direkt das globale Minimum, bei der hier vorgestellten Aufgabenstellung den Ort und die Form der Fehlstelle, zu bestimmen. Durch die Anwendung der genetischen Operatoren kann jedes lokale Minimum umgangen werden. Weiter benötigen beide Optimierungsmethoden keine Gradienteninformationen der jeweiligen Aufgabenstellung. Die aufwendige Berechnung des Differenzenquotienten wie bei Anwendung des SQP-Verfahrens oder des Kalman-Filters ent-

fallen hier. In dieser Arbeit wurde ein genetischer Algorithmus in der Weise umstrukturiert, daß eine Ganzzahloptimierung durchführbar wurde. Damit konnte zusätzlich zu den Geometriedaten der Defekte auch deren Anzahl bestimmt werden.

Der genetische Algorithmus weist allerdings nur eine langsame Konvergenzgeschwindigkeit auf. Die numerischen Tests zeigen nach wenigen Iterationsschritten eine gute Näherung der Parameter des zu suchenden Defektes, zu deren genauen Bestimmung wurden jedoch weitaus mehr Iterationen benötigt. Somit liegt der Nachteil des genetischen Algorithmus und der Evolutionsstrategie eindeutig in der zu hohen Rechenzeit. Um die Rechenzeiten zu senken, wurde im Zeitbereich eine parallelisierte Variante der Evolutionsstrategie getestet. Die Ergebnisse wurden nicht in dieser Arbeit aufgelistet. Die Parallelisierung ermöglicht zwar eine schnellere Konvergenz des Algorithmus, die Gesamtrechenzeiten bleiben aber weiterhin sehr hoch. Der Zeitbedarf für ausführlichere Studien ist mit den zur Zeit einsetzbaren Rechnern nicht vertretbar.

Ein neuronales Netz geht anders vor als die iterativen Optimierungsstrategien: Es muß ähnlich wie das menschliche Gehirn beim Lernvorgang mit Lerndaten angelernet werden. Erst nach erfolgtem Training kann das neuronale Netz eine Prognose einer zu bestimmenden Größe angeben.

Es muß somit ein erheblicher Aufwand betrieben werden, Lerndaten für das neuronale Netz zu generieren, d.h. es müssen für fast jede mögliche Kombination aus Lage und geometrischer Form der Schädigung Simulationen durchgeführt werden und dem neuronalen Netz als Eingangsdaten zur Verfügung gestellt werden. Der zeitliche Aufwand dieser Datenerstellung für die Identifikation von mehr als einem Defekt innerhalb einer Scheibe ist daher sehr hoch. Die Schadenserkennung mit einem neuronalen Netz wird deshalb auf eine einzelne Schädigung beschränkt.

Eine weitere Schwierigkeit besteht in der richtigen Auswahl der zur Verfügung stehenden Lerndaten. Beinhalten die Lerndaten zu viele Informationen, die keinen Aufschluß über die Lage der Fehlstellen geben, kann der Lernvorgang empfindlich gestört oder verhindert werden. Um diesen Effekt auszuschließen, wurden bei den Untersuchungen im Zeitbereich die berechneten Lerndaten gezielt reduziert. Es wurden nur Verschiebungen berücksichtigt, die in direktem Zusammenhang mit einer durch die Schädigung hervorgerufenen Wellenreflektion stehen. Schadenslokalisationen in der Statik ergaben gute Identifikationsergebnisse. Eine unbekannte Fehlstelle wurde grundsätzlich gut identifiziert.

Die Untersuchungen ergaben außerdem, daß der Einfluß des Trainingsverfahrens auf die Güte der Ergebnisse einen größeren Einfluß hat als die Netzarchitektur des Netzwerkes. Entgegen der Vermutung, daß ein „größeres“ Netz bessere Ergebnisse erzeugt, ist die Wahl eines passenden Lernverfahrens entscheidend. Aus den Untersuchungen ging das Levenberg-Marquardt Verfahren als geeignetste Lernstrategie hervor. Mit einem sehr kleinen Netz konnten sehr gute Ergebnisse erzielt werden, während bei Anwendung anderer Lernverfahren und größerer Netzarchitekturen keine brauchbaren Resultate erzielt werden konnten.

Nachteilig auf die Qualität der Ergebnisse wirkt sich die Anzahl der zu suchenden Defekte aus. Das neuronale Netz gibt auch bei einem nicht defekten Bauteil das Vorhandensein einer Schädigung an. Dieser Fehler konnte jedoch durch Anwendung mehrerer neuronaler Netze verhindert

werden. Das erste neuronale Netz kann das Bauteil als ungeschädigt oder als geschädigt einstufen. Durch das zweite Netz werden schließlich die Unbekannten der Schädigung bestimmt. Der entscheidende Vorteil der Methode des neuronalen Netze ist ihre praktische Anwendung: Ein einmal angelerntes Netz kann beliebig oft und sehr schnell eingesetzt werden. Das Aufwendige ist lediglich der Lernvorgang, der mit Simulationsdaten durchgeführt wird. Danach steht es ohne neue Berechnungen zur Erkennung von Schadensfällen zur Verfügung.

Eine hier erstmals entwickelte hybride Optimierungsstrategie kann ellipsoide Fehlstellen innerhalb einer dreidimensionalen Geometrie erfolgreich identifizieren. Dabei wird zunächst der genetische Algorithmus angewandt, anschließend wird die Genauigkeit der Prognose durch das Kalman-Filter verbessert.

Andere Optimierungsstrategien mit iterativem Charakter, wie zum Beispiel das Simulated Annealing, ähneln sehr stark dem genetischen Algorithmus. Die Untersuchungen mit dem genetischen Algorithmus haben deutlich gezeigt, daß eine ungefähre Lage der Schädigung recht schnell angegeben werden kann, während die Berechnung der exakten Position der Schädigung sehr zeitaufwendig ist. Aufgrund dessen wurde auf eine Untersuchung dieser Algorithmen verzichtet.

7 Zusammenfassung und Ausblick

Risse und Fehlstellen innerhalb hochbelasteter Bauteile sind die häufigste Ursache für tragische Unfälle im Bau- und Maschinenwesen. In dieser Arbeit werden numerische Methoden untersucht, um Risse innerhalb einfacher geometrischer Strukturen zu lokalisieren. Dabei wird nicht nur ein schadhaftes Bauteil als solches erkannt, sondern es wird die exakte Position und das Ausmaß der Schädigung identifiziert.

Zur Simulation des mechanischen Verhaltens des Bauteils wird die Randelementmethode verwendet, die im Vergleich zur Finiten-Elemente-Methode den Vorteil der einfachen Diskretisierung des Randes mit einer einfachen Implementierung einer Fehlstelle innerhalb der Geometrie verbindet. Da ein zerstörungsfreies Rißidentifizierungsverfahren getestet werden soll, beschränkt sich die Simulation auf das elastische Verformungsverhalten der untersuchten Geometrien.

Zur Lokalisierung der Schädigung werden in dieser Arbeit vier unterschiedliche Optimierungsstrategien zur Rißerkennung vorgestellt. Der genetische Algorithmus als das rechenaufwendigste Optimierungsalgorithmus ist als einziges Verfahren in der Lage globale Optima zu finden. Die Ergebnisse bestätigen die Robustheit und die Zuverlässigkeit des genetischen Algorithmus. Aufgrund des hohen Rechenaufwandes ist das Verfahren nur für die Rißidentifikation bei statischen Untersuchungen geeignet. Weiterhin wird das Kalman-Filter zur Defekterkennung herangezogen. Erste Untersuchungen zeigen, daß das Kalman-Filter schneller Fehlstellen identifizieren kann als die Sequential-Quadratic-Programming(SQP) Methode, die in dieser Arbeit zu Vergleichszwecken als Optimierungsstrategie verwendet. Im allgemeinen gilt das SQP-Verfahren als das zuverlässigste der Standardverfahren zur Optimierung von nichtlinearen Funktionen. Da dieses Verfahren sehr viele Iterationen benötigt, um ein akzeptables Ergebnis zu erhalten, ist es für diese Anwendung allerdings zu aufwendig.

Als weiteres Verfahren kann das neuronale Netz, das nicht unbedingt zu den Optimierungsstrategien gezählt wird, zur Rißerkennung verwendet werden. In dieser Arbeit werden ausschließlich neuronale Netze verwendet, die mit der Backpropagationregel angelernet werden. Unterschiedliche Netzwerkarchitekturen und Lernstrategien werden untersucht. Es zeigt sich, daß in Abhängigkeit des verwendeten Lernverfahrens die Identifikation erfolgreich beendet wird.

Durch die Anwendung des genetischen Algorithmus können lokale Minima überwunden werden. Unter alleiniger Anwendung werden gute Ergebnisse bei allerdings hohem Rechenaufwand erzielt. Eine erste gute Näherung der gesuchten Schadensparameter errechnet der genetische Algorithmus nach wenigen Iterationen. Er benötigt für eine akzeptable Genauigkeit aber

anschließend noch viele Optimierungsschritte. Aus diesem Grund wurde auf die Anwendung dieser Strategie bei den Untersuchungen in der Dynamik verzichtet.

Ein neu entwickeltes hybrides Verfahren, das den genetischen Algorithmus mit dem Kalman-Filter koppelt, kombiniert die Vorteile der beiden Verfahren. Die Untersuchungen zeigen, daß es im statischen Bereich das leistungsfähigste Verfahren zur Rißidentifikation darstellt. Der genetische Algorithmus berechnet innerhalb weniger Iterationen die ungefähre Lage der Fehlstelle. Anschließend berechnet das Kalman-Filter die exakten Daten der Fehlstelle.

Für dynamische Untersuchungen erweist sich das neuronale Netz als nützliches Verfahren. Es muß nur einmal trainiert werden und kann beliebig oft zur Fehlstellendiagnose verwendet werden.

Das Kalman-Filter gilt hier als das erfolgversprechendste Optimierungsverfahren zum Erkennen von Fehlstellen. Aus diesem Grund wurde die Verwendung des Kalman-Filters in dieser Arbeit detailliert untersucht.

Zusammenfassend kann festgestellt werden, daß die Identifikation unterschiedlicher Schädigungen mittels numerischer Verfahren erfolgreich durchgeführt werden kann. Die Belastungsart und die Anzahl der Lastfälle ist dabei auf die Geometrie des Prüfkörpers abzustimmen. Orts-, Lage- und Formparameter elliptischer und runder Schädigungen sowie kontaktbehaftete und nicht kontaktbehafteter Risse können erfolgreich ermittelt werden.

In Zukunft ist eine wesentliche Steigerung der Rechnerkapazität zu erwarten, so daß kompliziertere Bauteile untersucht werden können. Die Diskretisierung der Bauteile kann sehr fein vorgenommen werden, so daß selbst kritische Bauteilkerben detailliert simuliert werden können.

Mit steigender Rechnerleistung können Identifikationsvorgänge schadhafter Bauteile nahezu in Echtzeit durchgeführt werden. So könnte es möglich sein, ein schadhaftes Bauteil als solches zu erkennen, bevor es zu einem Ausfall kommt.

Literaturverzeichnis

- [1] Andrieux, S.: Reciprocity principle and crack identification. *Inverse Problems*, **15**(1), 59–66, 1999.
- [2] Antes, H.: *Anwendungen der Methode der Randelemente in der Elastodynamik*. B. G. Teubner, 1988.
- [3] Bayer, H.G.: *The Theory of Evolution Strategies*. Springer, 2001.
- [4] Bäck, T.; Fogel, D.B.; Michalewics, T.: *Evolutionary Computation 1*. Institute Of Physics Publishing, 2000.
- [5] Bäck, T.; Fogel, D.B.; Michalewics, T.: *Evolutionary Computation 2*. Institute Of Physics Publishing, 2000.
- [6] Brammer, K.; G.Stiffing: *Kalman-Bucy-Filter Deterministische Beobachtung und stochastische Filterung*. R. Oldenbourg Verlag, 1975.
- [7] Brammer, K.; Stiffing, G.: *Stochastische Grundlagen des Kalman-Bucy-Filter - Wahrscheinlichkeitsrechnung und Zufallsprozesse*. R. Oldenbourg Verlag, 1990.
- [8] Brown, R.G.; Hwang, P.Y.C.: *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley & Sons, 1997.
- [9] Bruehl, M.: Crack identification using electrostatic measurements. *RAIRO*, **35**(3), 2001.
- [10] Catlin, D.E.: *Estimation, Control, and the Discrete Kalman Filter*. Springer-Verlag, 1989.
- [11] Engelhardt, M.; Likas, A.; Stavroulakis, G.E.: Neural crack identification. In *Fifth World Congress on Computational Mechanics*. (Mang, H.A.; Rammerstorfer, F.G.; Eberhardsteiner, J.; Eberhardsteiner, J., Eds.), Austria, 2002, WCCM.
- [12] Engelhardt, M.; Stavroulakis, G.; Antes, H.: Schadenserkennung mit genetischen Algorithmen. ZAMM, 2002.
- [13] Gallego, R.; Dominguez, J.: Solving Transient Dynamic Crack Problems by the Hyper-singular Boundary Element Method. *Fatigue & Fracture of Engineering Materials & Structures*, **20**, 799–812, 1997.

- [14] Gallego, R.; Rus, G.: Identification of cracks and cavities using topological sensitivity boundary integral equation. *Computational Mechanics*, **33**, 154–163, 2004.
- [15] Geiger, C.; Kanzow, C.: *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer-Verlag, 2002.
- [16] Grewal, M.S.; Andrews, A.P.: *Kalman Filtering - Theory and Practice*. Prentice-Hall, 1993.
- [17] Guiggiani, M.; Krishnasamy, G.; Rudolphi, T.; Rizzo, F.: A General Algorithm for the Numerical Solution of Hypersingular Boundary Integral Equations. *ASME Journal of Applied Mechanics*, **59**, 604–614, 1992.
- [18] Hagan, M.T.; Demuth, H. B.; Beale, M.: *Neural Network Design*. PWS, 1996.
- [19] Hagan, M.T.; Menhaj, M.B.: Training Feedforward Networks with the Marquardt Algorithm. *IEEE transactions on neural networks*, **5**, 989–993, 1994.
- [20] He, Y.; Guo, D.; Chu, F.: Using genetic algorithms and finite element methods to detect shaft crack for rotor-bearing system. *Mathematics and Computers in Simulation*, **57**, 95–108, 2001.
- [21] Hwu, C.; Liang, Y.C.: Hole/crack identification by static strains from multiple loading modes. *AIAA Journal*, **39**(2), 315–324, 2001.
- [22] Isermann, R.: *Digitale Regelsysteme, Band 2*. Springer-Verlag, 1987.
- [23] Kohonen, T.: *Self-Organization and Associative Memory*. Springer-Verlag, 1987.
- [24] Krawczuk, M.: Spectral finite element and genetic algorithm for crack detection in cantilever rol. *Key Engineering Materials*, **204**, 241–250, 2001.
- [25] Krebs, V.: *Nichtlineare Filterung*. R. Oldenbourg Verlag, 1980.
- [26] Liang, Y.C.; Hwu, C.: On-line identification of holes/cracks in composite structures. *IOP Publishing*, **10**, 599–609, 2001.
- [27] Liu, G.R.; Han, X.; Lam, K.Y.: A combined genetic algorithm and nonlinear least squares method for material characterization using elastic waves (in press). *Computer Methods in Applied Mechanics and Engineering*, 2001.
- [28] Liu, S.W.; Huang, J.H.; Sung, J.C.; Lee, C.C.: Detection of cracks using neural networks and computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, **25**, 2831–2846, 2002.
- [29] Masters, T.: *Practical Neural Network Recipes in C++*. Academic Press, 1993.
- [30] Moller, M.F.: A scaled conjugate algorithm for fast supervised learning. *Neural Networks*, **6**, 525–533, 1993.

- [31] Murty, K.G.: *Linear Complementarity, Linear and Nonlinear Programming*. Heldermann Verlag Berlin, 1988.
- [32] Oishi, A.; Yamada, K.; Yoshimura, S.; Yagawa, G.; Nagai, S.; Matsuda, Y.: Neural network-based inverse analysis for defect identification with laser ultrasonics. *Research on Nondestructive Evaluation*, 79–95, 2001.
- [33] Partheymueller, P.: *Numerische Simulation der 3D-Rissausbreitung mit der Randelementmethode*. VDI, 1998.
- [34] Riedmiller, M.; Braun, H.: A direct adaptive method for faster backpropagation learning: The Rprop algorithm. *Proceedings of the IEEE International Conference on Neural Networks*. IEEE, 1993.
- [35] Rojas, R.: *Theorie der neuronalen Netze*. Springer-Verlag, 1996.
- [36] Rus, G.; Gallego, R.: Optimization algorithms for identification inverse problems with the boundary element method. *Engineering Analysis with Boundary Elements*, **26**(4), 315–327, 2002.
- [37] Schanz, M.: *Eine Randelementformulierung im Zeitbereich mit verallgemeinerten viskoelastischen Stoffgesetzen*. Dissertation, Institut für Mechanik, Universität der Bundeswehr Hamburg, 1994.
- [38] Schöneburg, E.: *Genetische Algorithmen und Evolutionsstrategien*. Addison-Wesley, 1996.
- [39] Spellucci, P.: *Numerische Verfahren der nichtlinearen Optimierung*. Birkhäuser Verlag, 1993.
- [40] Stavroulakis, G.E.: Impact-echo from a unilateral interlayer crack. LCP-BEM modelling and neural identification. *Engineering Fracture Mechanics*, **62**, 165–184, 1999.
- [41] Stavroulakis, G.E.: *Inverse and Crack Identification Problems in Engineering Mechanics (Applied Optimization)*. Kluwer Academic Publishers, Dordrecht, Boston, London, 2001. Habilitationsschrift, Fakultät für Bauingenieurwesen, Technische Universität Carolo Wilhelmina zu Braunschweig,.
- [42] Stavroulakis, G.E.; Antes, H.: Flaw identification in elastomechanics: BEM simulation with local and genetic optimization. *Structural Optimization*, **16**(2/3), 162–175, 1998.
- [43] Stavroulakis, G.E.; Antes, H.: Neural crack identification in steady state elastodynamics. *Computer Methods in Applied Mechanics and Engineering*, **165**(1/4), 129–146, 1998.
- [44] Stavroulakis, G.E.; Engelhardt, M.; R.Gallego: Crack identification in elasticity. In *Fifth World Congress on Computational Mechanics*. (Mang, H.A.; Rammerstorfer, F.G.; Eberhardsteiner, J.; Eberhardsteiner, J., Eds.), Austria, 2002, WCCM.

- [45] Suh, M.-W.: Crack identification using hybrid neuro-genetic technique. *Journal of Sound and Vibration*, **238**(4), 617–636, 2000.
- [46] Tanaka, M.; Matsumoto, T.; Yamamura, H.: Application of BEM with extended Kalman filter to parameter identification of an elastic plate under dynamic loading. *Engineering Analysis with Boundary Elements*, **28**, 213–219, 2004.
- [47] Weikl, W.; Andrae, H.; Schnack, E.: An alternating iterative algorithm for the reconstruction of internal cracks in a three-dimensional body. *Inverse Problems*, **17**(6), 1957–1975, 2001.
- [48] Yoshimura, S.; Saito, Y.; Yagawa, G.: Identification of Two Dissimilar Surface Cracks Hidden in Solid Using Neural Networks and Computational Mechanics. *Computer Modeling and Simulation in Engineering*, 477–491, 1996.
- [49] Zang, C.; Imregun, M.: FRF-based structural damage detection using Kohonen self-organising maps. *International Journal of Acoustics and Vibration*, **5**(4), 167–172, 2000.
- [50] Zang, C.; Imregun, M.: Structural damage detection using artificial neural networks and measured FRF data reduced via principal component projection. *Journal of Sound and Vibration*, **242**(5), 813–827, 2001.
- [51] Zarchan, P.; Musoff, H.: *Fundamentals of Kalman Filtering: A Practical Approach*. American Institute of Aeronautics and Astronautics, Inc., 2000.
- [52] Zell, A.: *Simualtion Neuronaler Netze*. Addison-Wesley, 1996.

Braunschweiger Schriften zur Mechanik – BSM

Bisher erschienene Berichte in dieser Reihe

- 1–1990 Plonski, Thomas:
Dynamische Analyse von schnelldrehenden Kreiszylinderschalen
- 2–1991 Wegener, Konrad:
Zur Berechnung grosser plastischer Deformationen mit einem Stoffgesetz vom Über-
spannungstyp
- 3–1992 Gröhlich, Hubert:
Finite-Element-Formulierung für vereinheitlichte inelastische Werkstoffmodelle oh-
ne explizite Fliessflächenformulierung
- 4–1992 Hesselbarth, Hanfried:
Simulation von Versetzungsstrukturbildung, Rekristallisation und Kriechschädigung
mit dem Prinzip der zellulären Automaten
- 5–1992 Schlums, Hartmut:
Ein stochastisches Werkstoffmodell zur Beschreibung von Kriechen und zyklischem
Verhalten metallischer Werkstoffe
- 6–1992 Kublik, Frithjof:
Vergleich zweier Werkstoffmodelle bei ein- und mehrachsigen Versuchsführungen
im Hochtemperaturbereich
- 7–1992 Bechtloff, Jürgen:
Interpolationsverfahren höheren Grades für Robotersteuerungen
- 8–1993 Müller, Michael:
Dreidimensionale elastodynamische Analyse von Tanks mit fluidbenetzten Einbau-
ten
- 9–1993 Senker, Peter:
Stabilitätsanalyse elastischer RotorsystemeStabilitätsanalyse elastischer Rotorsyste-
me
- 10–1993 Cheng, Weimin:
Schallabstrahlung einer schwingenden Reisner/Mindlin Platte
- 11–1993 Wiebe, Thomas:
Wellenausbreitung in poroelastischen Medien: Untersuchung mit Randintegralglei-
chungen

- 12–1993 Hahne, Matthias:
Beschreibung der plastischen Längsdehnung bei Torsion mit einem makroskopischen Stoffgesetz
- 13–1993 Heisig, Gerald:
Zum statischen und dynamischen Verhalten von Tiefbohrsträngen in räumlich gekrümmten Bohrlöchern
- 14–1994 de Araújo, Francisco Célio:
Zeitbereichslösung linearer dreidimensionaler Probleme der Elastodynamik mit einer gekoppelten BE/FE-Methode
- 15–1994 Kristen, Martin:
Untersuchungen zur elektrischen Ansteuerung von Formgedächtnis-Antrieben in der Handhabungstechnik
- 16–1994 Latz, Kersten:
Dynamische Interaktion von Flüssigkeitsbehältern und Baugrund
- 17–1994 Jäger, Monika:
Entwicklung eines effizienten Randelementverfahrens für bewegte Schallquellen
- 18–1994 August, Martin:
Schwingungen und Stabilität eines elastischen Rades, das auf einer nachgiebigen Schiene rollt
- 19–1995 Erbe, Matthias:
Zur Simulation von Risswachstum in dreidimensionalen, elastisch-plastischen Strukturen mit der Methode der Finiten Elemente
- 20–1995 Gerdes, Ralf:
Ein stochastisches Werkstoffmodell für das inelastische Materialverhalten metallischer Werkstoffe im Hoch- und Tieftemperaturbereich
- 21–1995 Tröndle, Georg:
Effiziente Schallberechnung mit einem adaptiven Mehrgitterverfahren für die 3-D Randelementmethode
- 22–1996 Degenhardt, Richard:
Nichtlineare dynamische Bauwerksprobleme und Interaktion mit dem Baugrund
- 23–1996 Feise, Hermann Josef:
Modellierung des mechanischen Verhaltens von Schüttgütern
- 24–1996 Haubrok, Dietmar:
Reibungsfreie Kontaktprobleme der 2-D Elastostatik und -dynamik als Optimierungsaufgabe mit REM-Matrizen

- 25–1996 Lehmann, Lutz:
Numerische Simulation der Spannungs- und Geschwindigkeitsfelder in Silos mit Einbauten
- 26–1996 Klein, Ralf:
Dynamische Interaktion von dünnwandigen Tragwerken und Boden mit Abschirmschlitzten
- 27–1996 Kopp, Thilo:
Simulation grosser inelastischer Deformationen bei Torsionsversuchen
- 28–1997 Harder, Jörn:
Simulation lokaler Fliessvorgänge in Polykristallen
- 29–1997 Lewerenz, Malte Christian:
Zur numerischen Behandlung von Werkstoffmodellen für zeitabhängig plastisches Materialverhalten
- 30–1997 Meywerk, Martin:
Stabilität und Verschleiss bei auf Schienen laufenden Eisenbahnradstätzen
- 31–1997 Plagge, Frank:
Nichtlineares, inelastisches Verhalten von Spiralseilen
- 32–1997 Neubert, Michael:
Richtungsregelung beim Tiefbohren
- 33–1998 Sangi, Daryoush:
Die Versetzungsstrukturbildung in Metallen
- 34–1998 Thielecke, Frank:
Parameteridentifizierung von Simulationsmodellen für das viskoplastische Verhalten von Metallen - Theorie, Numerik, Anwendung -
- 35–1998 Vietgen, Jürgen:
Numerische Simulation duktilen Risswachstums unter Berücksichtigung von Schädigung
- 36–1998 Lär, Bernard:
Einfluss transienter Anregungen auf die Zylinderkopf-Akustik
- 37–1998 Scheld, Christian:
Auswirkungen dynamischer Interaktionen auf das Schwingungsverhalten von Tanks
- 38–1999 Baaran, Jens:
Schallfeldanalyse bei sich bewegenden schallerzeugenden Körpern

- 39–1999 Daros, Carlos:
Wellenausbreitung in unendlich ausgedehnten piezoelektrischen Medien mit transversal isotroper Symmetrie
- 40–2000 Nils, Wagner:
Untersuchung der Boden-Fahrzeug-Interaktion mit gekoppelten Rand- und Finite-Element-Methoden
- 41–2001 Langer, Sabine:
Schalltransmission durch Isolierverglasung
- 42–2001 Schacht, T.:
Orientierungsabhängige Rissbildung in duktilen Metallen
- 43–2002 Ackerman, Lutz:
Simulation der Schalltransmission durch Wände
- 44–2002 Barthold, Franz-Joseph:
Zur Kontinuumsmechanik inverser Geometrie Probleme
- 45–2002 Böhrnsen, Jens-Uwe:
Dynamisches Verhalten von Schüttgütern beim Entleeren aus Silos
- 46–2002 Zorn, C.:
Plastisch instabile Verformung aufgrund dynamischer Reckalterung und korrelierten Versetzungsgleitens
- 47–2002 Küsel, M.:
Wellige Verschleissmuster auf Laufflächen von Eisenbahnrädern
- 48–2002 Bross, S.:
Versetzungsdynamik und Reckalterung als Ursache instabilen Materialverhaltens beim Portevin-Le Chatelier-Effekt
- 49–2002 Vesper, Matthias:
Modellierung von Korn und Korngrenze in polykristallinen Gefügen
- 50–2003 Steindorf, Jan:
Partitionierte Verfahren für Probleme der Fluid-Struktur Wechselwirkung
- 51–2003 Meyer, Marcus:
Reduktionsmethoden zur Simulation des aeroelastischen Verhaltens von Windkraftanlagen
- 52–2003 Wittich, Hauke:
Inhomogene Gleitung bei nahgeordneten CuAl-Legierungen

- 53–2003 Hupfer, Knut:
Einfluss der Mikrostruktur auf die Festigkeitseigenschaften von Metallschäumen
- 54–2004 Schmelzer, Martin:
Identifikation der Parameter von Zeitbereichsmodellen linear-viskoelastischer Werkstoffe
- 55–2004 Ostendorf, Michael:
Geschwindigkeitsmessungen in Silos mit der Particle Image Velocimetry